Cisco Nexus 7000 Series NX-OS Multicast Routing Configuration Guide

First Published: 2016-12-23

Last Modified: 2020-03-24

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED “AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NONINFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

All printed copies and duplicate soft copies of this document are considered uncontrolled. See the current online version for the latest version.

Cisco has more than 200 offices worldwide. Addresses and phone numbers are listed on the Cisco website at www.cisco.com/go/offices.

The documentation set for this product strives to use bias-free language. For purposes of this documentation set, bias-free is defined as language that does not imply discrimination based on age, disability, gender, racial identity, ethnic identity, sexual orientation, socioeconomic status, and intersectionality. Exceptions may be present in the documentation due to language that is hardcoded in the user interfaces of the product software, language used based on standards documentation, or language that is used by a referenced third-party product.

CONTENTS

CHAPTER 7

CHAPTER 8

Preface

This preface describes the audience, organization, and conventions of the Book Title. It also provides information on how to obtain related documentation.

This chapter includes the following topics:

Audience

This publication is for experienced network administrators who configure and maintain Cisco NX-OS on Cisco Nexus 7000 Series Platform switches.

Document Conventions

Note

- As part of our constant endeavor to remodel our documents to meet our customers' requirements, we have modified the manner in which we document configuration tasks. As a result of this, you may find a deviation in the style used to describe these tasks, with the newly included sections of the document following the new format.
- The Guidelines and Limitations section contains general guidelines and limitations that are applicable to all the features, and the feature-specific guidelines and limitations that are applicable only to the corresponding feature.

Command descriptions use the following conventions:

Convention	Description
bold	Bold text indicates the commands and keywords that you enter literally as shown.
Italic	Italic text indicates arguments for which the user supplies the values.
[x]	Square brackets enclose an optional element (keyword or argument).
[x | y]	Square brackets enclosing keywords or arguments separated by a vertical bar indicate an optional choice.
{x | y}	Braces enclosing keywords or arguments separated by a vertical bar indicate a required choice.
[x {y | z}]	Nested set of square brackets or braces indicate optional or required choices within optional or required elements. Braces and a vertical bar within square brackets indicate a required choice within an optional element.
`variable`	Indicates a variable for which you supply values, in context where italics cannot be used.

Convention	Description
string	A nonquoted set of characters. Do not use quotation marks around the string or the string will include the quotation marks.

Examples use the following conventions:

Convention	Description
`screen font`	Terminal sessions and information the switch displays are in screen font.
`boldface screen font`	Information you must enter is in boldface screen font.
italic screen font	Arguments for which you supply values are in italic screen font.
< >	Nonprinting characters, such as passwords, are in angle brackets.
[ ]	Default responses to system prompts are in square brackets.
!, #	An exclamation point (!) or a pound sign (#) at the beginning of a line of code indicates a comment line.

This document uses the following conventions:

Note Means *reader take note*. Notes contain helpful suggestions or references to material not covered in the manual.

Caution Means *reader be careful*. In this situation, you might do something that could result in equipment damage or loss of data.

Related Documentation

Documentation for Cisco Nexus 7000 Series Switches is available at:

- Configuration Guides

  http://www.cisco.com/c/en/us/support/switches/nexus-7000-series-switches/products-installation-and-configuration-guides-list.html

- Command Reference Guides

  http://www.cisco.com/c/en/us/support/switches/nexus-7000-series-switches/products-command-reference-list.html

- Release Notes

  http://www.cisco.com/c/en/us/support/switches/nexus-7000-series-switches/products-release-notes-list.html

- Install and Upgrade Guides

  http://www.cisco.com/c/en/us/support/switches/nexus-7000-series-switches/products-installation-guides-list.html

- Licensing Guide

  http://www.cisco.com/c/en/us/support/switches/nexus-7000-series-switches/products-licensing-information-listing.html

Documentation for Cisco Nexus 7000 Series Switches and Cisco Nexus 2000 Series Fabric Extenders is available at the following URL:

http://www.cisco.com/c/en/us/support/switches/nexus-2000-series-fabric-extenders/products-installation-and-configuration-guides-list.html

Documentation Feedback

To provide technical feedback on this document, or to report an error or omission, please send your comments to nexus7k-docfeedback@cisco.com. We appreciate your feedback.

Communications, Services, and Additional Information

- To receive timely, relevant information from Cisco, sign up at Cisco Profile Manager.
- To get the business impact you're looking for with the technologies that matter, visit Cisco Services.
- To submit a service request, visit Cisco Support.
- To discover and browse secure, validated enterprise-class apps, products, solutions and services, visit Cisco Marketplace.
- To obtain general networking, training, and certification titles, visit Cisco Press.
- To find warranty information for a specific product or product family, access Cisco Warranty Finder.

Cisco Bug Search Tool

Cisco Bug Search Tool (BST) is a web-based tool that acts as a gateway to the Cisco bug tracking system that maintains a comprehensive list of defects and vulnerabilities in Cisco products and software. BST provides you with detailed defect information about your products and software.

CHAPTER 1

New and Changed Information

This chapter describes new and changed features.

New and Changed Information

Your software release might not support all the features in this document. For the latest caveats and feature information, see the Bug Search Tool at https://tools.cisco.com/bugsearch/ and the release notes for your software release.

Table 1: New and Changed Information for Multicast Routing

Feature	Description	Changed in Release	Where Documented
PIM Allow RP (IPv6)	This feature is introduced.	8.4(2)	Information about PIM Allow RP, on page 127Enabling PIM Allow RP, on page 130
PIM Allow RP	This feature is introduced.	8.4(1)	Information about PIM Allow RP, on page 127Enabling PIM Allow RP, on page 130
Configuring MoFRR	This feature is introduced.	8.2(1)	Configuring MoFRR, on page 185
Configuring Multicast Extranet	This feature is introduced.	8.2(1)	Configuring Multicast Extranet, on page 179

CHAPTER 2

Overview

This chapter describes the multicast features of Cisco NX-OS.

Licensing Requirements

For a complete explanation of Cisco NX-OS licensing recommendations and how to obtain and apply licenses, see the *Cisco NX-OS Licensing Guide*.

Information About Multicast

IP multicast is a method of forwarding the same set of IP packets to a number of hosts within a network. You can use multicast in both IPv4 and IPv6 networks to provide efficient delivery of data to multiple destinations.

Note Beginning with Cisco NX-OS Release 5.2(1) for the Nexus 7000 Series devices, you can configure Protocol-Independent Multicast v4 (PIMv4) to run over generic routing encapsulation (GRE) tunnels including outgoing interfaces (OIF). In prior Cisco NX-OS releases, tunnel interfaces do not support PIM.

Note Beginning with Cisco NX-OS Release 7.3(0)DX(1), multicast generic routing encapsulation (mGRE) tunnels is supported on M3 Series modules.

Multicast involves both a method of delivery and discovery of senders and receivers of multicast data, which is transmitted on IP multicast addresses called groups. A multicast address that includes a group and source IP address is often referred to as a channel. The Internet Assigned Number Authority (IANA) has assigned

224.0.0.0 through 239.255.255.255 as IPv4 multicast addresses. For more information, see http://www.iana.org/assignments/multicast-addresses.

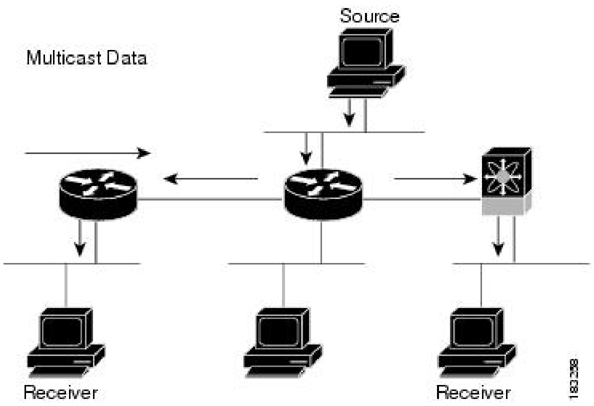
IPv6 multicast addresses begin with 0xFF. The IPv6 addressing architecture is defined by RFC 4291. For more information about the IANA reserved addresses, see http://www.iana.org/assignments/ipv6-multicast-addresses.

Note For a complete list of RFCs related to multicast, see IETF RFCs for IP Multicast, on page 203.

The routers in the network listen for receivers to advertise their interest in receiving multicast data from selected groups. The routers then replicate and forward the data from sources to the interested receivers. Multicast data for a group is transmitted only to those LAN segments with receivers that requested it.

This figure shows one source transmitting multicast data that is delivered to two receivers. In the figure, because the center host is on a LAN segment where no receiver requested multicast data, no data is delivered to that receiver.

Figure 1: Multicast Traffic from One Source to Two Receivers



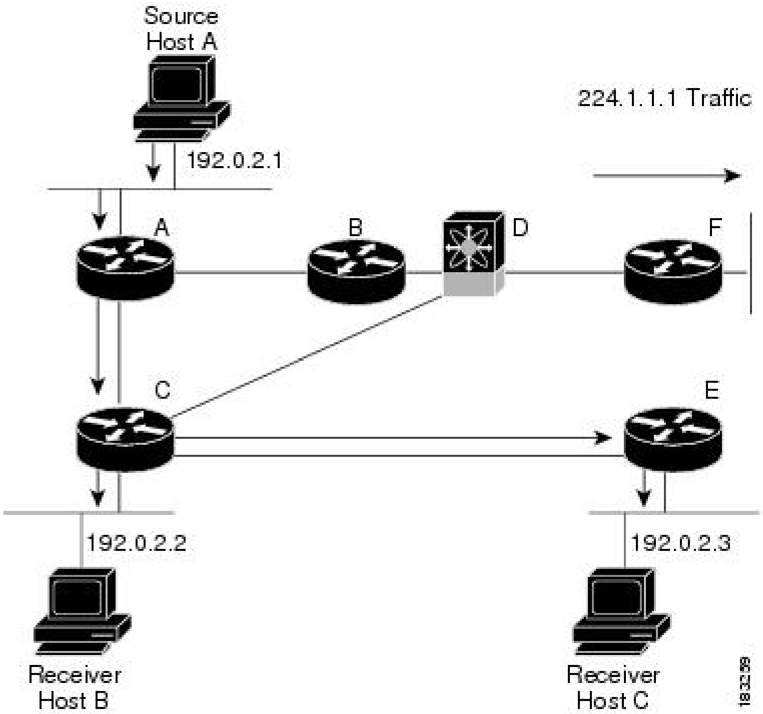
Multicast Distribution Trees

A multicast distribution tree represents the path that multicast data takes between the routers that connect sources and receivers. The multicast software builds different types of trees to support different multicast methods.

Source Trees

A source tree represents the shortest path that the multicast traffic takes through the network from the sources that transmit to a particular multicast group to receivers that requested traffic from that same group. Because of the shortest path characteristic of a source tree, this tree is often referred to as a shortest path tree (SPT). This figure shows a source tree for group 224.1.1.1 that begins at host A and connects to hosts B and C.

Figure 2: Source Tree



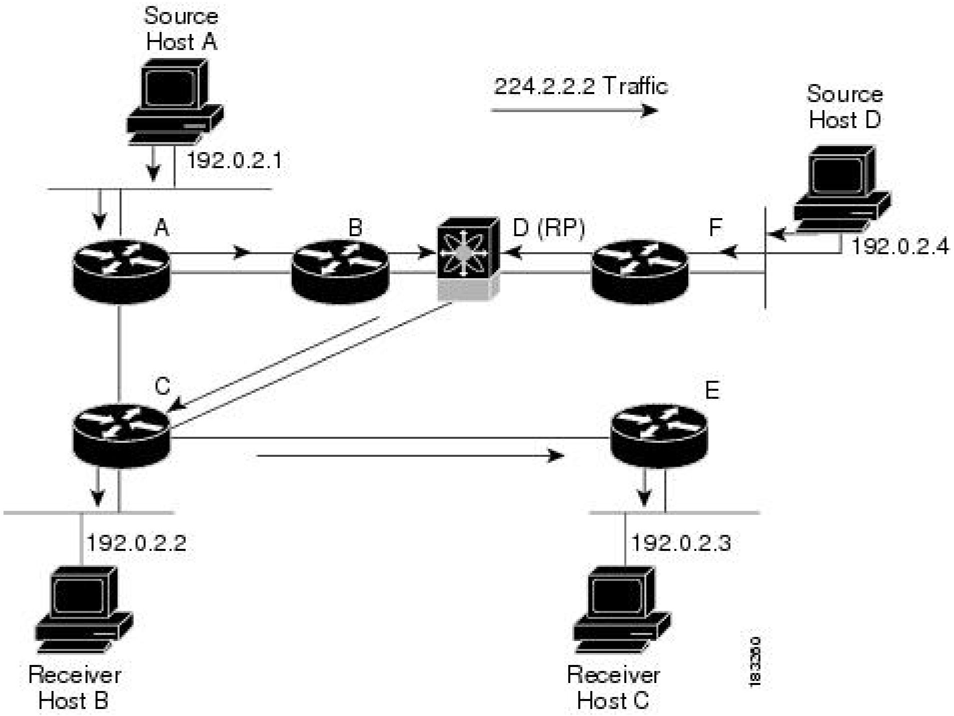
The notation (S, G) represents the multicast traffic from source S on group G. The SPT in this figure is written (192.0.2.1, 224.1.1.1). Multiple sources can be transmitting on the same group.

Shared Trees

A shared tree represents the shared distribution path that the multicast traffic takes through the network from a shared root or rendezvous point (RP) to each receiver. (The RP creates an SPT to each source.) A shared tree is also called an RP tree (RPT). The figure below shows a shared tree for group 224.1.1.1 with the RP at router D. Source hosts A and D send their data to router D, the RP, which then forwards the traffic to receiver hosts B and C.

The notation (*, G) represents the multicast traffic from any source on group G. The shared tree in this figure is written (*, 224.2.2.2).

Figure 3: Shared Tree

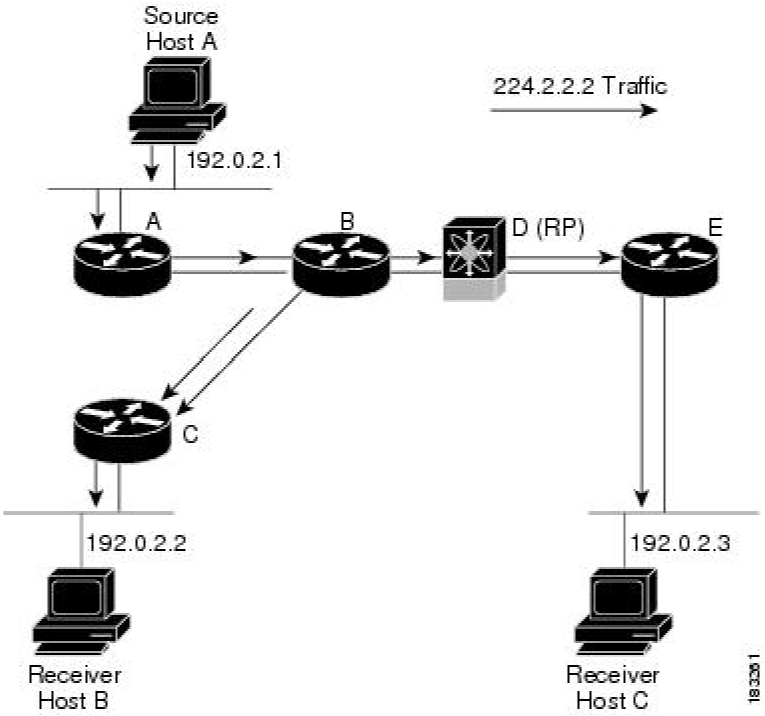


Bidirectional Shared Trees

A bidirectional shared tree represents the shared distribution path that the multicast traffic takes through the network from a shared root, or rendezvous point (RP), to each receiver. Multicast data is forwarded to receivers encountered on the way to the RP. The advantage of the bidirectional shared tree is shown in the figure below. Multicast traffic flows directly from host A to host B through routers B and C. In a shared tree, the data from source host A is first sent to the RP (router D) and then forwarded to router B for delivery to host B.

The notation (*, G) represents the multicast traffic from any source on group G. The bidirectional tree in the figure below is written (*, 224.2.2.2).

Figure 4: Bidirectional Shared Tree



Multicast Forwarding

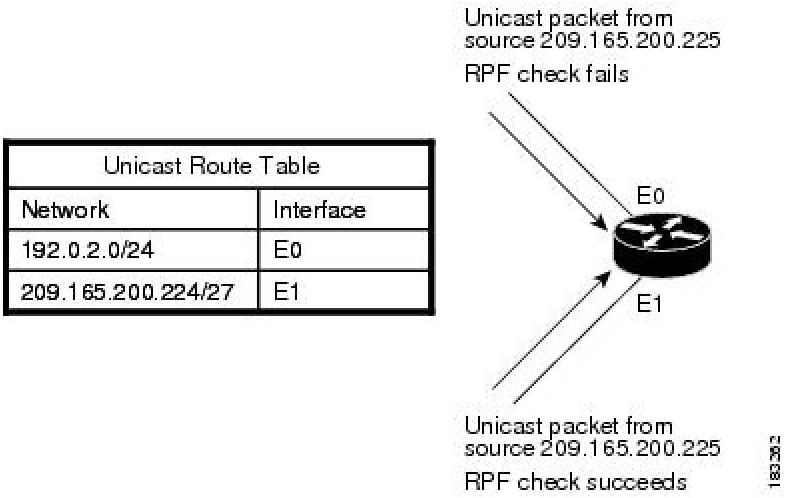
Because multicast traffic is destined for an arbitrary group of hosts, the router uses reverse path forwarding (RPF) to route data to active receivers for the group. When receivers join a group, a path is formed either toward the source (SSM mode) or the RP (ASM or Bidir mode). The path from a source to a receiver flows in the reverse direction from the path that was created when the receiver joined the group.

For each incoming multicast packet, the router performs an RPF check. If the packet arrives on the interface leading to the source, the packet is forwarded out each interface in the outgoing interface (OIF) list for the group. Otherwise, the router drops the packet.

Note In Bidir mode, if a packet arrives on a non-RPF interface, and the interface was elected as the designated forwarder (DF), then the packet is also forwarded in the upstream direction toward the RP.

The figure below shows an example of RPF checks on packets coming in from different interfaces. The packet that arrives on E0 fails the RPF check because the unicast route table lists the source of the network on interface E1. The packet that arrives on E1 passes the RPF check because the unicast route table lists the source of that network on interface E1.

Figure 5: RPF Check Example



Unicast Route Table	
Network	Interface
192.0.2.0/24	E0
209.165.200.224/27	E1

Cisco NX-OS PIM and PIM6

Cisco NX-OS supports multicasting with Protocol Independent Multicast (PIM) sparse mode. PIM is IP routing protocol independent and can leverage whichever unicast routing protocols are used to populate the unicast routing table. In PIM sparse mode, multicast traffic is sent only to locations of the network that specifically request it. PIM dense mode is not supported by Cisco NX-OS.

Note In this publication, the term “PIM” is used for PIM sparse mode version 2.

To access multicast commands, you must enable the PIM or PIM6 feature. Multicast is enabled only after you enable PIM or PIM6 on an interface of each router in a domain. You configure PIM for an IPv4 network and PIM6 for an IPv6 network. By default, IGMP and MLD are running on the system.

PIM, which is used between multicast-capable routers, advertises group membership across a routing domain by constructing multicast distribution trees. PIM builds shared distribution trees on which packets from multiple sources are forwarded, as well as source distribution trees, on which packets from a single source are forwarded.

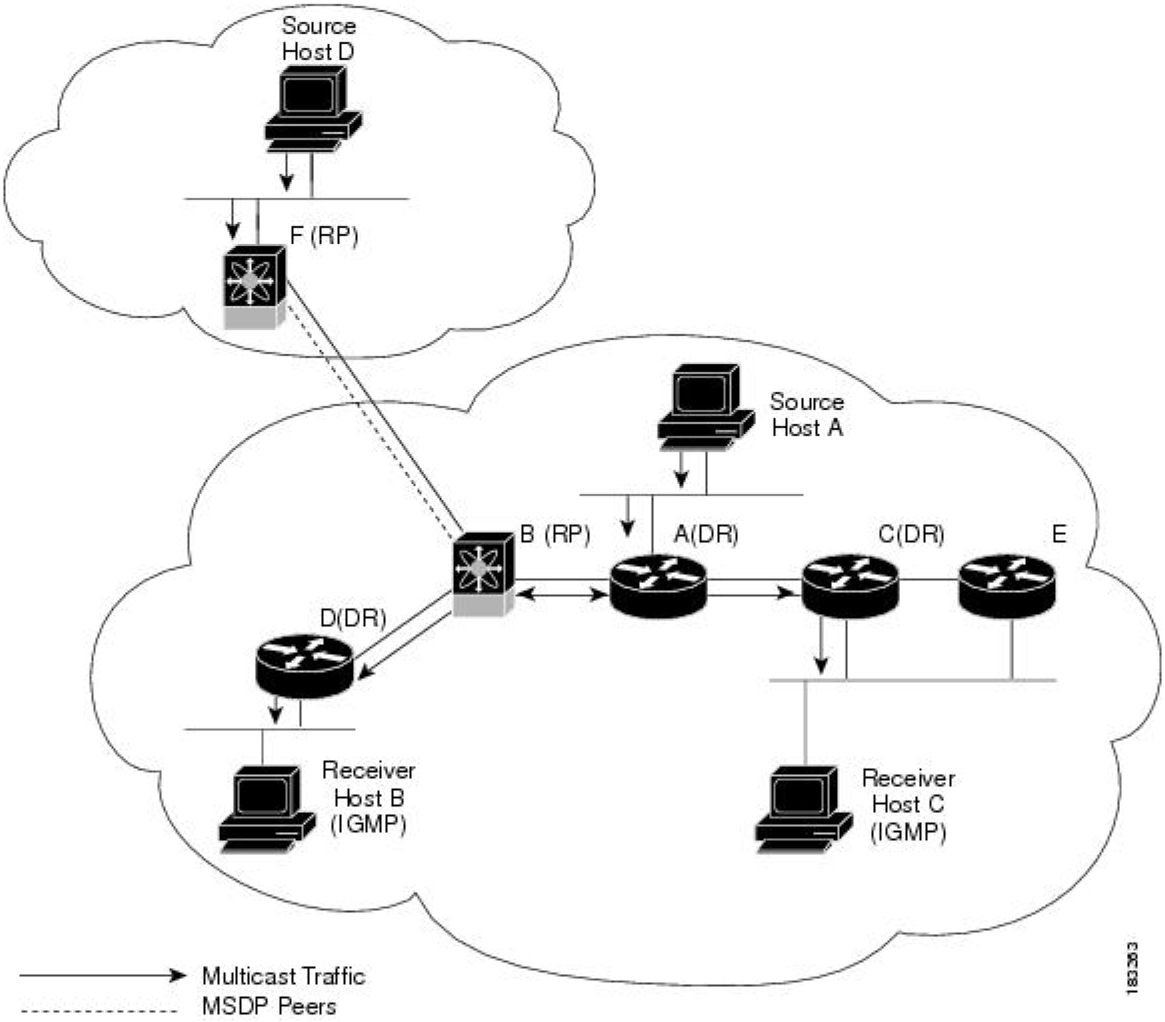
The distribution trees change automatically to reflect the topology changes due to link or router failures. PIM dynamically tracks both multicast-capable sources and receivers, although the source state is not created in Bidir mode.

The router uses the unicast routing table and RPF routes for multicast to create multicast routing information. In Bidir mode, additional routing information is created.

Note In this publication, “PIM for IPv4” and “PIM6 for IPv6” refer to the Cisco NX-OS implementation of PIM sparse mode. A PIM domain can include both an IPv4 and an IPv6 network.

The figure below shows two PIM domains in an IPv4 network.

Figure 6: PIM Domains in an IPv4 Network



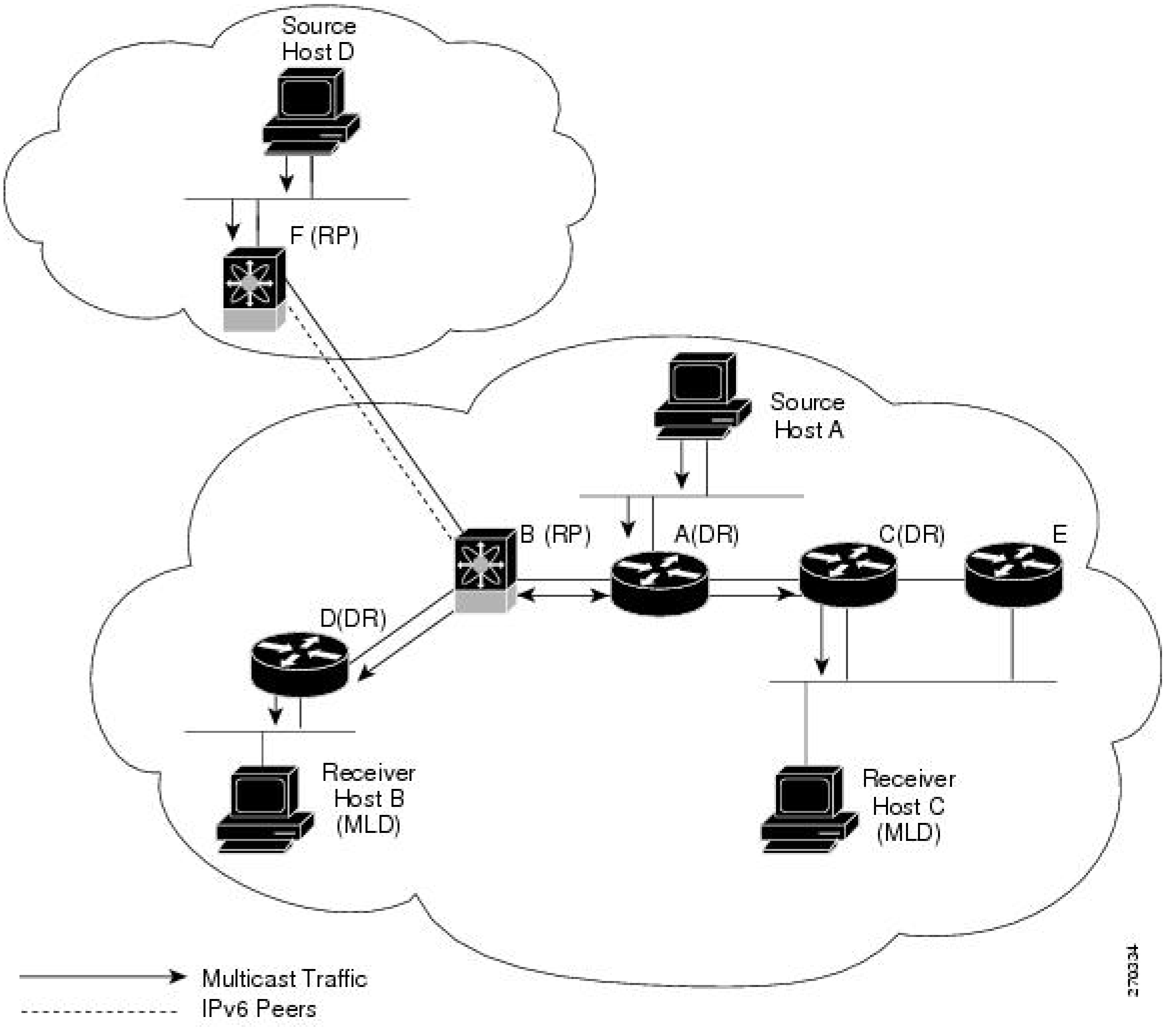
- The lines with arrows show the path of the multicast data through the network. The multicast data originates from the sources at hosts A and D.
- The dashed line connects routers B and F, which are Multicast Source Discovery Protocol (MSDP) peers. MSDP supports the discovery of multicast sources in other PIM domains.
- Hosts B and C receive multicast data by using the Internet Group Management Protocol (IGMP) to advertise requests to join a multicast group.
- Routers A, C, and D are designated routers (DRs). When more than one router is connected to a LAN segment, such as C and E, the PIM software chooses one router to be the DR so that only one router is responsible for putting multicast data on the segment

Router B is the rendezvous point (RP) for one PIM domain and router F is the RP for the other PIM domain. The RP provides a common point for connecting sources and receivers within a PIM domain.

This figure shows two PIM6 domains in an IPv6 network. In an IPv6 network, receivers that want to receive multicast data use the Multicast Listener Discovery (MLD) protocol to advertise requests to join a multicast group. MSDP, which allows for discovery of multicast sources in other PIM domains, is not supported for

IPv6. You can configure IPv6 peers and use Source-Specific Multicast (SSM) and Multiprotocol BGP (MBGP) to forward multicast data between PIM6 domains.

Figure 7: PIM6 Domains in an IPv6 Network



PIM supports three multicast modes for connecting sources and receivers:

- Any source multicast (ASM)
- Source-specific multicast (SSM)
- Bidirectional shared trees (Bidir)

Cisco NX-OS supports a combination of these modes for different ranges of multicast groups. You can also define RPF routes for multicast.

ASM

Any Source Multicast (ASM) is a PIM tree building mode that uses shared trees to discover new sources and receivers as well as source trees to form shortest paths from receivers to sources. The shared tree uses a network node as the root, called the rendezvous point (RP). The source tree is rooted at first-hop routers, directly attached to each source that is an active sender. The ASM mode requires an RP for a group range. An RP can be configured statically or learned dynamically by the Auto-RP or BSR group-to-RP discovery protocols. If an RP is learned and is not known to be a Bidir-RP, the group operates in ASM mode.

The ASM mode is the default mode when you configure RPs.

Bidir

Bidirectional shared trees (Bidir) is a PIM mode that, like the ASM mode, builds a shared tree between receivers and the RP, but does not support switching over to a source tree when a new receiver is added to a group. In the Bidir mode, the router that is connected to a receiver is called the designated forwarder because multicast data can be forwarded directly from the designated router (DR) to the receiver without first going to the RP. The Bidir mode requires that you configure an RP.

The Bidir mode can reduce the amount of resources required on a router when there are many multicast sources and can continue to operate whether or not the RP is operational or connected.

SSM

Source-Specific Multicast (SSM) is a PIM mode that builds a source tree that originates at the designated router on the LAN segment that receives a request to join a multicast source. Source trees are built by sending PIM join messages in the direction of the source. The SSM mode does not require you to configure RPs.

The SSM mode allows receivers to connect to sources outside the PIM domain.

RPF Routes for Multicast

You can configure static multicast RPF routes to override what the unicast routing table uses. This feature is used when the multicast topology is different than the unicast topology.

IGMP and MLD

By default, the Internet Group Management Protocol (IGMP) for PIM and Multicast Listener Discovery (MLD) for PIM6 are running on the system.

IGMP and MLD protocols are used by hosts that want to receive multicast data to request membership in multicast groups. Once the group membership is established, multicast data for the group is directed to the LAN segment of the requesting host.

You can configure IGMPv2 or IGMPv3 on an interface. You will usually configure IGMPv3 to support SSM mode. By default, the software enables IGMPv2.

You can configure MLDv1 or MLDv2 on an interface. You will usually configure MLDv2 to support SSM mode. By default, the software enables MLDv2.

IGMP Snooping

IGMP snooping is a feature that limits multicast traffic on VLANs to the subset of ports that have known receivers. By examining (snooping) IGMP membership report messages from interested hosts, multicast traffic is sent only to VLAN ports that interested hosts reside on. By default, IGMP snooping is running on the system.

Interdomain Multicast

Cisco NX-OS provides several methods that allow multicast traffic to flow between PIM domains.

SSM

The PIM software uses SSM to construct a shortest path tree from the designated router for the receiver to a known source IP address, which may be in another PIM domain. The ASM and Bidir modes cannot access sources from another PIM domain without the use of another protocol.

Once you enable PIM or PIM6 in your networks, you can use SSM to reach any multicast source that has an IP address known to the designated router for the receiver.

MSDP

Multicast Source Discovery Protocol (MSDP) is a multicast routing protocol that is used with PIM to support the discovery of multicast sources in different PIM domains.

Note Cisco NX-OS supports the PIM Anycast-RP, which does not require MSDP configuration.

MBGP

Multiprotocol BGP (MBGP) defines extensions to BGP4 that enable routers to carry multicast routing information. PIM and PIM6 can use this multicast information to reach sources in external BGP autonomous systems.

For information about MBGP, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Command Reference*.

MRIB and M6RIB

The Cisco NX-OS IPv4 Multicast Routing Information Base (MRIB) is a repository for route information that is generated by multicast protocols such as PIM and IGMP. The MRIB does not affect the route information itself. The MRIB maintains independent route information for each virtual routing and forwarding (VRF) instance in a virtual device context (VDC). For more information about VDCs, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*.

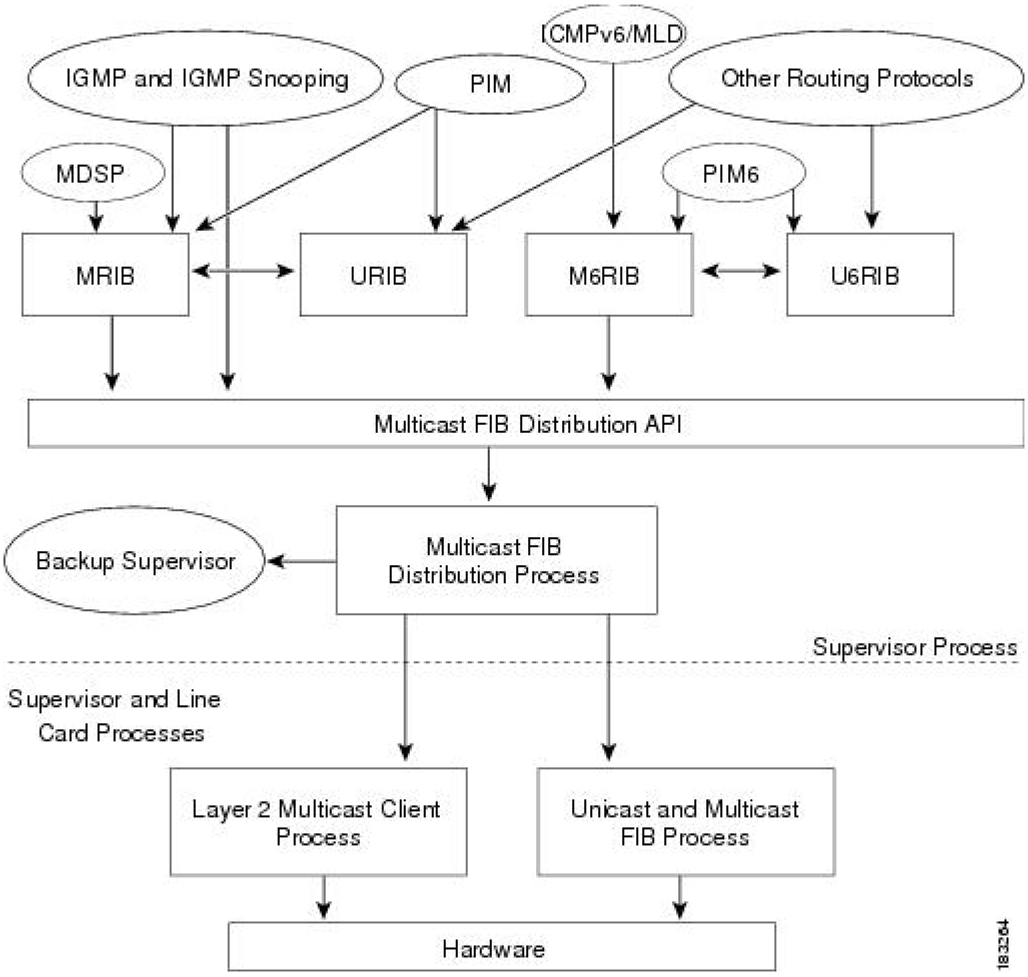
Similar to the MRIB for IPv4 routing information, the M6RIB maintains IPv6 routing information that is generated by protocols such as PIM6 and MLD.

This figure shows the major components of the Cisco NX-OS multicast software architecture:

- The Multicast FIB (MFIB and M6FIB) Distribution (MFDM) API defines an interface between the multicast Layer 2 and Layer 3 control plane modules, including the MRIB and M6RIB, and the platform forwarding plane. The control plane modules send the Layer 3 route update and Layer 2 lookup information using the MFDM API.
- The multicast FIB distribution process distributes the multicast update messages to all the relevant modules and the standby supervisor. It runs only on the supervisor.
- The Layer 2 multicast client process sets up the Layer 2 multicast hardware forwarding path. It runs on both the supervisor and the modules.

- The unicast and multicast FIB process manages the Layer 3 hardware forwarding path. It runs on both the supervisor and the modules.

Figure 8: Cisco NX-OS Multicast Software Architecture


ICMPv6/MLD
IGMP and IGMP Snooping
PIM
Other Routing Protocols
MDSP
PIM6
MRIB
URIB
M6RIB
U6RIB
Multicast FIB Distribution API
Backup Supervisor
Multicast FIB Distribution Process
Supervisor Process
Supervisor and Line Card Processes
Layer 2 Multicast Client Process
Unicast and Multicast FIB Process
Hardware
183264

MRIB/M6RIB Dynamic Shared Memory Support

The Cisco NX-OS IPv4 Multicast Routing Information Base and IPv6 Multicast Routing Information Base (MRIB/M6RIB) dynamic shared memory support feature supports dynamic shared memory in a virtual device context (VDC). The MRIB/M6RIB dynamic shared memory feature changes the shared memory dynamically based on the number of routes that are added or removed from the MRIB/M6RIB. Instead of a static allocation of the entire configured memory for the multicast routes, the shared memory for MRIB/M6RIB dynamically adds up or is removed based on the increase or decrease, respectively, in the number of routes.

This feature also ensures that information on the shared memory is accessible and readable by the MRIB/M6RIB clients during a dynamic change in the shared memory. The MRIB/M6RIB dynamic shared memory feature also supports device switchover (from active to standy state and vice-versa) when the shared memory increases or decreases.

Dynamic Shared Memory support in MRIB/M6RIB for VDC

The MRIB and M6RIB maintain independent route information for each virtual routing and forwarding (VRF) instance in a virtual device context (VDC). VDC resource templates set the minimum and maximum limits for the shared memory when you create a VDC. The Cisco NX-OS software reserves the minimum limit for the resource to the VDC. Any resources allocated to the VDC beyond the minimum are based on the maximum limit and availability on the device. VDC templates set limits on both IPv4 multicast route memory and IPv6 multicast route memory. You can change the VDC resource limits by applying a new VDC resource template. Changes to the limits take effect immediately except for the IPv4 and IPv6 route memory limits, which take effect after the next VDC reset, physical device reload, or physical device stateful switchover. A switchover occurs when the active route processor (RP) fails, is removed from the networking device, or is manually taken down for maintenance.

Instead of a static allocation of the entire configured memory for the multicast routes, the shared memory for MRIB/M6RIB dynamically adds up or is removed based on the increase or decrease, respectively, in the number of routes, without making any modifications to the VDC.

The dynamic shared memory in MRIB/M6RIB is not affected during synchronization of the active and standby processors and during a physical device stateful switchover from the active to the standby processor.

Virtual Port Channels and Multicast

A virtual port channel (vPC) allows a single device to use a port channel across two upstream switches. When you configure a vPC, the following multicast features may be affected:

- PIM and PIM6—Cisco NX-OS software for the Nexus 7000 Series devices does not support PIM SSM or Bidr on a vPC.
- GMP snooping—You should configure the vPC peers identically.

Maximum Transmission Unit Limitation

On the Cisco NX-OS software for the Nexus 7000 Series devices, the Maximum Transmission Unit (MTU) for a given mroute is equal to the smallest MTU of the OIF. Packets exceeding that MTU value are dropped and not multicast routed to any of the OIFs for that mroute.

Multicasting with both F Series and M Series Modules in a Chassis

Beginning with Cisco NX-OS Release 5.1, you can add an F Series module, which is a Layer 2-only module, into the Cisco Nexus 7000 Series chassis. When you add this module to a chassis that already contains M Series modules, you can provision multicasting.

General Multicast Restrictions

Cisco NX-OS multicast features have the following restrictions:

- Cisco Nexus 7000 Series devices do not support Pragmatic General Multicast (PGM).

High-Availability Requirements for Multicast

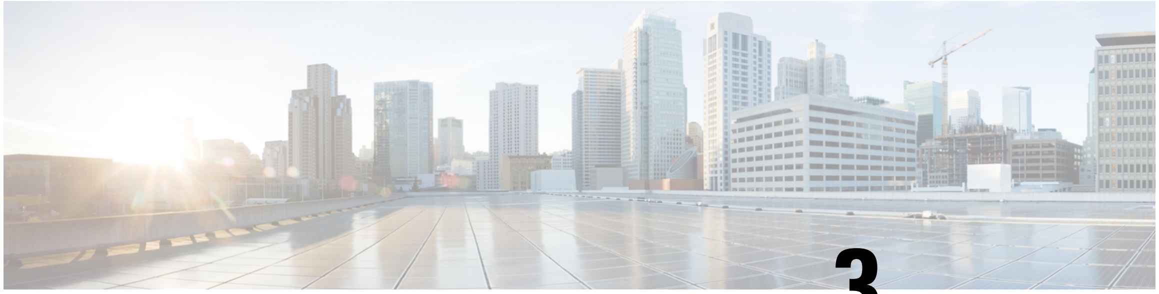
After a multicast routing protocol is restarted, its state is recovered from the MRIB process. When a supervisor switchover occurs, the MRIB recovers its state from the hardware, and the multicast protocols recover their state from periodic message activity. For more information about high availability, see the *Cisco Nexus 7000 Series NX-OS High Availability and Redundancy Guide*.

Related Documents

Related Topic	Document Title
VDCs	*Cisco Nexus 7000 Series NX-OS Virtual Device Context Command Reference*
CLI Commands	*Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*

Technical Assistance

Description	Link
Technical Assistance Center (TAC) home page, containing 30,000 pages of searchable technical content, including links to products, technologies, solutions, technical tips, and tools. Registered Cisco.com users can log in from this page to access even more content.	https://www.cisco.com/c/en/us/support/index.html



CHAPTER 3

Configuring IGMP

This chapter describes how to configure the Internet Group Management Protocol (IGMP) on Cisco NX-OS devices for IPv4 networks.

Information About IGMP

IGMP is an IPv4 protocol that a host uses to request multicast data for a particular group. Using the information obtained through IGMP, the software maintains a list of multicast group or channel memberships on a per-interface basis. The systems that receive these IGMP packets send multicast data that they receive for requested groups or channels out the network segment of the known receivers.

By default, the IGMP process is running. You cannot enable IGMP manually on an interface. IGMP is automatically enabled when you perform one of the following configuration tasks on an interface:

- Enable PIM
- Statically bind a local multicast group
- Enable link-local group reports

IGMP Versions

The device supports IGMPv2 and IGMPv3, as well as IGMPv1 report reception.

By default, the software enables IGMPv2 when it starts the IGMP process. You can enable IGMPv3 on interfaces where you want its capabilities.

IGMPv3 includes the following key changes from IGMPv2:

- Support for Source-Specific Multicast (SSM), which builds shortest path trees from each receiver to the source, through the following features:
  - Host messages that can specify both the group and the source.
  - The multicast state that is maintained for groups and sources, not just for groups as in IGMPv2.

- Hosts no longer perform report suppression, which means that hosts always send IGMP membership reports when an IGMP query message is received.

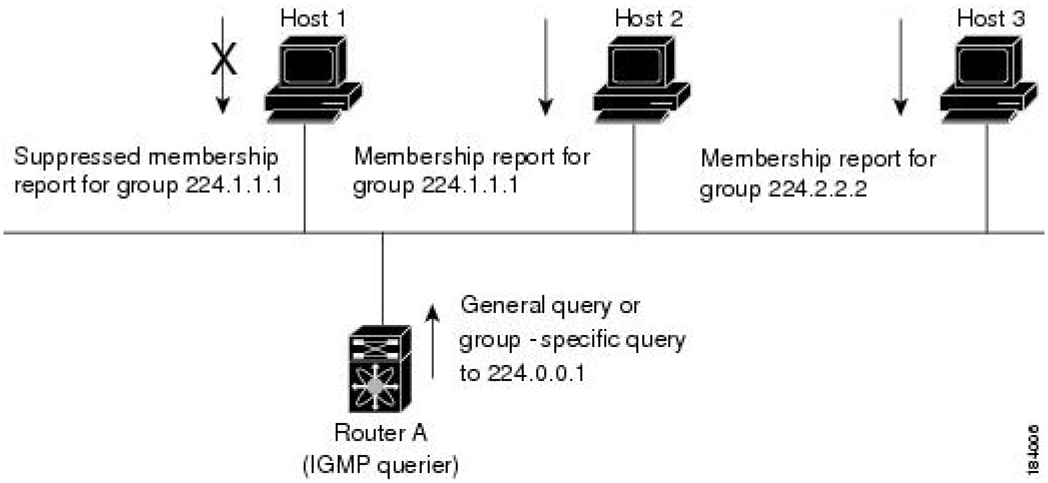
For detailed information about IGMPv2, see RFC 2236.

For detailed information about IGMPv3, see RFC 3376.

IGMP Basics

The basic IGMP process of a router that discovers multicast hosts is shown in the figure below. Hosts 1, 2, and 3 send unsolicited IGMP membership report messages to initiate receiving multicast data for a group or channel.

Figure 9: IGMPv1 and IGMPv2 Query-Response Process


Host 1
Host 2
Host 3
Suppressed membership report for group 224.1.1.1
Membership report for group 224.1.1.1
Membership report for group 224.2.2.2
General query or group - specific query to 224.0.0.1
Router A (IGMP querier)
184006

In the figure below, router A, which is the IGMP designated querier on the subnet, sends query messages to the all-hosts multicast group at 224.0.0.1 periodically to discover whether any hosts want to receive multicast data. You can configure the group membership timeout value that the router uses to determine that no members of a group or source exist on the subnet. For more information about configuring the IGMP parameters, see *Configuring IGMP Interface Parameters*.

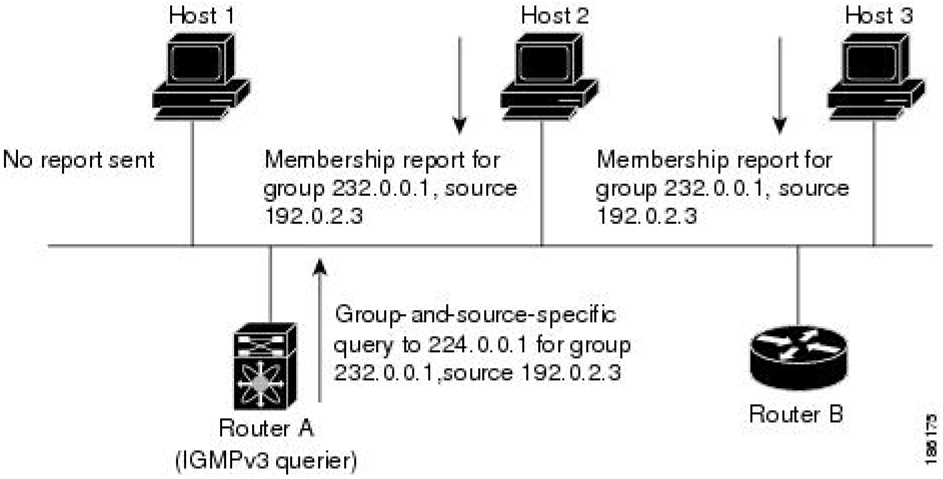
The software elects a router as the IGMP querier on a subnet if it has the lowest IP address. As long as a router continues to receive query messages from a router with a lower IP address, it resets a timer that is based on its querier timeout value. If the querier timer of a router expires, it becomes the designated querier. If that router later receives a host query message from a router with a lower IP address, it drops its role as the designated querier and sets its querier timer again.

In this figure, host 1's membership report is suppressed and host 2 sends its membership report for group 224.1.1.1 first. Host 1 receives the report from host 2. Because only one membership report per group needs to be sent to the router, other hosts suppress their reports to reduce network traffic. Each host waits for a random time interval to avoid sending reports at the same time. You can configure the query maximum response time parameter to control the interval in which hosts randomize their responses.

Note IGMPv1 and IGMPv2 membership report suppression occurs only on hosts that are connected to the same port.

In this figure, router A sends the IGMPv3 group-and-source-specific query to the LAN. Hosts 2 and 3 respond to the query with membership reports that indicate that they want to receive data from the advertised group and source. This IGMPv3 feature supports SSM. For information about configuring SSM translation to support SSM for IGMPv1 and IGMPv2 hosts, see *Configuring an IGMP SSM Translation*.

Figure 10: IGMPv3 Group-and-Source-Specific Query


Host 1
Host 2
Host 3
No report sent
Membership report for group 232.0.0.1, source 192.0.2.3
Membership report for group 232. 0. 0. 1, source 192. 0. 2. 3
Group- and-source-specific query to 224. 0. 0. 1 for group 232. 0. 0. 1,source 192. 0. 2. 3
Router A (IGMPv3 querier)
Router B
186173

Note IGMPv3 hosts do not perform IGMP membership report suppression.

Messages sent by the designated querier have a time-to-live (TTL) value of 1, which means that the messages are not forwarded by the directly connected routers on the subnet. You can configure the frequency and number of query messages sent specifically for IGMP startup, and you can configure a short query interval at startup so that the group state is established as quickly as possible. Although usually unnecessary, you can tune the query interval used after startup to a value that balances the responsiveness to host group membership messages and the traffic created on the network.

Caution Changing the query interval can severely impact multicast forwarding.

When a multicast host leaves a group, a host that runs IGMPv2 or later sends an IGMP leave message. To check if this host is the last host to leave the group, the software sends an IGMP query message and starts a timer that you can configure called the last member query response interval. If no reports are received before the timer expires, the software removes the group state. The router continues to send multicast traffic for a group until its state is removed.

You can configure a robustness value to compensate for packet loss on a congested network. The robustness value is used by the IGMP software to determine the number of times to send messages.

Link local addresses in the range 224.0.0.0/24 are reserved by the Internet Assigned Numbers Authority (IANA). Network protocols on a local network segment use these addresses; routers do not forward these addresses because they have a TTL of 1. By default, the IGMP process sends membership reports only for nonlink local addresses, but you can configure the software to send reports for link local addresses.

For more information about configuring the IGMP parameters, see *Configuring IGMP Interface Parameters*.

Virtualization Support

A virtual device context (VDC) is a logical representation of a set of system resources. Within each VDC, you can define multiple virtual routing and forwarding (VRF) instances. One IGMP process can run per VDC. The IGMP process supports all VRFs in that VDC and performs the function of IGMP snooping within that VDC. For information about IGMP snooping, see *Configuring IGMP Snooping*.

You can use the **show** commands with a VRF argument to provide a context for the information displayed. The default VRF is used if no VRF argument is supplied.

For information about configuring VDCs, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*.

For information about configuring VRFs, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

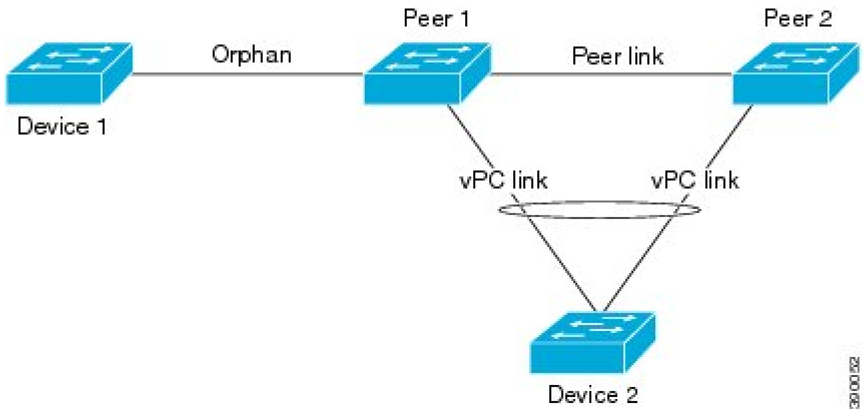
IGMP vPC Incremental Sync

The IGMP vPC incremental sync feature enables routes on the virtual port channel (vPC) peer to synchronize with other routes while the peer link is being established. This feature is a Layer 2 IPv4 multicast feature that enables faster convergence in vPC topologies. This feature enables Layer 2 Internet Group Management Protocol (IGMP) states to be synchronized between vPC peer devices in a triggered and incremental manner instead of periodic synchronization.

Overview of IGMP vPC Incremental Sync

The IGMP vPC Incremental Sync feature sends incremental updates to the peer link using Cisco Fabric Service (CFS), instead of sending all Join and Leave messages. The routes between peers are synced while the peer link is being set up.

Figure 11: Sample topology for implementing IGMP vPC Incremental Sync



Peer 1 is a vPC peer that receives the join/query/protocol independent multicast (PIM) hello either from Device 1 or from Device 2, which is on the vPC link. Peer 2 is a vPC peer that receives incremental updates from Peer 1 on the CFS. Device 1 acts as an orphan. Any port that is not configured as a vPC, but carries a vPC VLAN, is called an orphan.

The vPC peer link synchronizes states between the vPC peer devices. In addition to carrying control traffic between two VPC devices, the vPC peer link also carries multicast and broadcast data traffic. In some link failure scenarios, it also carries unicast traffic.

Interfaces that receive Query and PIM hello are added as device ports. Interfaces that receive Join messages are added as group outgoing interfaces (OIFs). Interfaces that receive Leave messages, delete the OIF from the group entry.

Benefits of IGMP vPC Incremental Sync

- Reduces CFS congestion.
- Results in faster convergence.

Prerequisites for IGMP vPC Incremental Sync

vPC peers must have the same version of the Cisco software image.

Verifying IGMP vPC Incremental Sync

Command	Purpose
show ip igmp internal vpc	Displays the summary of the IGMP vPC incremental sync configuration.

Prerequisites for IGMP

IGMP has the following prerequisites:

- You are logged onto the device.
- You are in the correct virtual device context (VDC). A VDC is a logical representation of a set of system resources. You can use the **switchto vdc** command with a VDC number.
- For global configuration commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.

Default Settings for IGMP

This table lists the default settings for IGMP parameters.

Table 2: Default IGMP Parameters

Parameters	Default
IGMP version	2
Startup query interval	30 seconds
Startup query count	2
Robustness value	2
Querier timeout	255 seconds
Query timeout	255 seconds

Parameters	Default
Query max response time	10 seconds
Query interval	125 seconds
Last member query response interval	1 second
Last member query count	2
Group membership timeout	260 seconds
Report link local multicast groups	Disabled
Enforce router alert	Disabled
Immediate leave	Disabled

Configuring IGMP Parameters

You can configure the IGMP global and interface parameters to affect the operation of the IGMP process.

Note If you are familiar with the Cisco IOS CLI, be aware that the Cisco NX-OS commands for this feature might differ from the Cisco IOS commands that you would use.

Configuring IGMP Interface Parameters

You can configure the optional IGMP interface parameters described in this table.

Table 3: IGMP Interface Parameters

Parameter	Description
IGMP version	IGMP version that is enabled on the interface. The IGMP version can be 2 or 3. The default is 2.

Parameter	Description
Static multicast groups	Multicast groups that are statically bound to the interface. You can configure the groups to join the interface with the (*, G) state or specify a source IP to join with the (S, G) state. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. **Note** Although you can configure the (S, G) state, the source tree is built only if you enable IGMPv3. For information about SSM translation, see *Configuring an IGMP SSM Translation*. You can configure a multicast group on all the multicast-capable routers on the network so that pinging the group causes all the routers to respond.
Static multicast groups on OIF	Multicast groups that are statically bound to the output interface. You can configure the groups to join the output interface with the (*, G) state or specify a source IP to join with the (S, G) state. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. **Note** Although you can configure the (S, G) state, the source tree is built only if you enable IGMPv3. For information about SSM translation, see the *Configuring an IGMP SSM Translation*.
Startup query interval	Startup query interval. By default, this interval is shorter than the query interval so that the software can establish the group state as quickly as possible. Values range from 1 to 18,000 seconds. The default is 31 seconds.
Startup query count	Number of queries sent at startup that are separated by the startup query interval. Values range from 1 to 10. The default is 2.
Robustness value	Robustness variable that you can tune to reflect expected packet loss on a congested network. You can increase the robustness variable to increase the number of times that packets are resent. Values range from 1 to 7. The default is 2.

Parameter	Description
Querier timeout	Number of seconds that the software waits after the previous querier has stopped querying and before it takes over as the querier. Values range from 1 to 65,535 seconds. The default is 255 seconds.
Query max response time	Maximum response time advertised in IGMP queries. You can tune the burstiness of IGMP messages on the network by setting a larger value so that host responses are spread out over a longer time. This value must be less than the query interval. Values range from 1 to 25 seconds. The default is 10 seconds.
Query interval	Frequency at which the software sends IGMP host query messages. You can tune the number of IGMP messages on the network by setting a larger value so that the software sends IGMP queries less often. Values range from 1 to 18,000 seconds. The default is 125 seconds.
Last member query response interval	Interval in which the software sends a response to an IGMP query after receiving a host leave message from the last known active host on the subnet. If no reports are received in the interval, the group state is deleted. You can use this value to tune how quickly the software stops transmitting on the subnet. The software can detect the loss of the last member of a group or source more quickly when the values are smaller. Values range from 1 to 25 seconds. The default is 1 second.
Last member query count	Number of times that the software sends an IGMP query, separated by the last member query response interval, in response to a host leave message from the last known active host on the subnet. Values range from 1 to 5. The default is 2. Setting this value to 1 means that a missed packet in either direction causes the software to remove the multicast state from the queried group or channel. The software may wait until the next query interval before the group is added again.
Group membership timeout	Group membership interval that must pass before the router decides that no members of a group or source exist on the network. Values range from 3 to 65,535 seconds. The default is 260 seconds.
Report link local multicast groups	Option that enables sending reports for groups in 224.0.0.0/24. Link local addresses are used only by protocols on the local network. Reports are always sent for nonlink local groups. The default is disabled.

Parameter	Description
Report policy	Access policy for IGMP reports that is based on a route-map policy. [1]
Access groups	Option that configures a route-map policy to control the multicast groups that hosts on the subnet serviced by an interface can join. **Note** Only the **match ip multicast group** command is supported in this route map policy. The **match ip address** command for matching an ACL is not supported.
Immediate leave	Option that minimizes the leave latency of IGMPv2 group memberships on a given IGMP interface because the device does not send group-specific queries. When immediate leave is enabled, the device will remove the group entry from the multicast routing table immediately upon receiving a leave message for the group. The default is disabled. **Note** Use this command only when there is one receiver behind the interface for a given group.

[1] To configure route-map policies, see the Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide.

For information about configuring multicast route maps, see *Configuring Route Maps to Control RP Information Distribution*.

Procedure

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td><b>Step 1</b></td><td><b>config t</b>
<b>Example:</b>
<code>switch# config t</code>
<code>switch(config)#</code></td><td>Enters configuration mode.</td></tr>
<tr><td><b>Step 2</b></td><td><b>interface</b> <i>interface</i>
<b>Example:</b>
<code>switch(config)# interface ethernet 2/1</code>
<code>switch(config-if)#</code></td><td>Enters interface mode on the interface type and number, such as <i>ethernet slot/port.</i></td></tr>
<tr><td><b>Step 3</b></td><td><table><tr><th>Option</th><th>Description</th></tr><tr><td><b>ip igmp version</b> <i>value</i></td><td>Sets the IGMP version to the value specified.</td></tr></table></td><td>These commands are used to configure the IGMP interface parameters.</td></tr>
</table>

	Command or Action	Purpose
	Option	**Description**
	`switch(config-if)# ip igmp version 3`	Values can be 2 or 3. The default is 2. The **no** form of the command sets the version to 2.

	ip igmp join-group {**group** [**source** *source*] | **route-map** *policy-name*}
	`switch(config-if)# ip igmp join-group 230.0.0.0`
		Statically binds a multicast group to the outgoing interface, which is handled by the device hardware. If you specify only the group address, the (*, G) state is created. If you specify the source address, the (S, G) state is created. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command.
		Note A source tree is built for the (S, G) state only if you enable IGMPv3.

	Command or Action	Purpose

Option	Description
	Caution The device CPU must be able to handle the traffic generated by using this command. Because of CPU load constraints, using this command, especially in any form of scale, is not recommended. Consider using the **ip igmp static-oif** command instead.

ip igmp static-oif {*group* [**source** *source*] | **route-map** *policy-name*}
`switch(config-if)# ip igmp static-oif 230.0.0.0`
	Statically binds a multicast group to the outgoing interface, which is handled by the device hardware. If you specify only the group address, the (*, G) state is created. If you specify the source address, the (S, G) state is created. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command.
	Note A source tree is built for the (S, G) state only if you enable IGMPv3.

ip igmp startup-query-interval *seconds*	Sets the query interval used when the software

	Command or Action	Purpose
	Option	**Description**
	switch(config-if)# ip igmp startup-query-interval 25	starts up. Values can range from 1 to 18,000 seconds. The default is 31 seconds.
	ip igmp startup-query-count *count* switch(config-if)# ip igmp startup-query-count 3	Sets the query count used when the software starts up. Values can range from 1 to 10. The default is 2.
	ip igmp robustness-variable *value* switch(config-if)# ip igmp robustness-variable 3	Sets the robustness variable. You can use a larger value for a lossy network. Values can range from 1 to 7. The default is 2.
	ip igmp querier-timeout *seconds* switch(config-if)# ip igmp querier-timeout 300	Sets the querier timeout that the software uses when deciding to take over as the querier. Values can range from 1 to 65,535 seconds. The default is 255 seconds.
	ip igmp query-timeout *seconds* switch(config-if)# ip igmp query-timeout 300	Sets the query timeout that the software uses when deciding to take over as the querier. Values can range from 1 to 65,535 seconds. The default is 255 seconds. **Note** This command has the same functionality as the **ip igmp querier-timeout** command.
	ip igmp query-max-response-time *seconds* Example switch(config-if)# ip igmp query-max-response-time 15	Sets the response time advertised in IGMP queries. Values can range from 1 to 25 seconds. The default is 10 seconds.
	ip igmp query-interval *interval*	Sets the frequency at which the software sends

	Command or Action	Purpose

Option	Description
`switch(config-if)# ip igmp query-interval 100`	IGMP host query messages. Values can range from 1 to 18,000 seconds. The default is 125 seconds.
ip igmp last-member-query-response-time *seconds* `switch(config-if)# ip igmp last-member-query-response-time 3`	Sets the query interval waited after sending membership reports before the software deletes the group state. Values can range from 1 to 25 seconds. The default is 1 second.
ip igmp last-member-query-count *count* `switch(config-if)# ip igmp last-member-query-count 3`	Sets the number of times that the software sends an IGMP query in response to a host leave message. Values can range from 1 to 5. The default is 2.
ip igmp group-timeout *seconds* `switch(config-if)# ip igmp group-timeout 300`	Sets the group membership timeout for IGMPv2. Values can range from 3 to 65,535 seconds. The default is 260 seconds.
ip igmp report-link-local-groups `switch(config-if)# ip igmp report-link-local-groups`	Enables sending reports for groups in 224.0.0.0/24. Reports are always sent for nonlink local groups. By default, reports are not sent for link local groups.
ip igmp report-policy *policy* `switch(config-if)# ip igmp report-policy my_report_policy`	Configures an access policy for IGMP reports that is based on a route-map policy.
ip igmp access-group *policy* `switch(config-if)# ip igmp access-group my_access_policy`	Configures a route-map policy to control the multicast groups that hosts on the subnet serviced by an interface can join.

	Command or Action	Purpose
	Option	**Description**
		Note Only the **match ip multicast group** command is supported in this route map policy. The **match ip address** command for matching an ACL is not supported.

	`ip igmp immediate-leave`
	`switch(config-if)# ip igmp immediate-leave`
		Enables the device to remove the group entry from the multicast routing table immediately upon receiving a leave message for the group. Use this command to minimize the leave latency of IGMPv2 group memberships on a given IGMP interface because the device does not send group-specific queries. The default is disabled.
		Note Use this command only when there is one receiver behind the interface for a given group.

Step 4	**show ip igmp interface** [*interface*] [**vrf** *vrf-name* | **all**] [**brief**] **Example:** `switch(config)# show ip igmp interface`	(Optional) Displays IGMP information about the interface.
Step 5	**copy running-config startup-config** **Example:**	(Optional) Saves configuration changes.

	Command or Action	Purpose
	`switch(config)# copy running-config startup-config`	

Configuring an IGMP SSM Translation

You can configure an SSM translation to provide SSM support when the router receives IGMPv1 or IGMPv2 membership reports. Only IMPv3 provides the capability to specify group and source addresses in membership reports. By default, the group prefix range is 232.0.0./8. To modify the PIM SSM range, see *Configuring SSM*.

The Internet Group Management Protocol (IGMP) Source-Specific Multicast (SSM) Translation feature enables a SSM-based multicast core network to be deployed when the multicast host do not support IGMPv3 or is forced to send group joins instead of (S,G) reports to interoperate with layer-2 switches. The IGMP SSM-Translation feature provides the functionality to configure multiple sources for the same SSM group. Protocol Independent Multicast (PIM) must be configured on the device before configuring the SSM translation.

This Table lists the example SSM Translations.

Table 4: Table 3 Example SSM Translation

group Prefix	Source Address
232.0.0.0/8	10.1.1.1
232.0.0.0/8	10.2.2.2
232.1.0.0/16	10.3.3.3
232.1.1.0/24	10.4.4.4

This Table shows the resulting MRIB routes that the IGMP process creates when it applies an SSM translation to the IMP membership report. If more than one translation applies, the router creates the (S,G) state for each translation.

Table 5: Table 4 Example Result of Applying SSM Translations

IGMPv2 membership Report	Resulting MRIB Route
232.1.1.1	(10.4.4.4, 232.1.1.1)
232.2.2.2	(10.1.1.1, 232.2.2.2)(10.2.2.2, 232.2.2.2)

Note This feature is similar to SSM mapping found in some Cisco IOS software.

The SSM translation configures source addresses per Virtual Routing and Forwarding (VRF) mode on the device to be mapped to specific SSM group ranges received in an IGMP report. The MRIB creates the (S,G) state rather than (*, G) state.

The IGMP SSM-Translation works in the following way:

- When an IGMPv1 or IGMPv2 report is received on an interface, the IGMP querier performs a translation table search for the reporting group.
- If there are configured source entries for the reporting group, the IGMP process adds to the interface that the report is received on to an (Si,G) entry corresponding to each configured source Si. These entries are stored in the MRIB for software and hardware multicast forwarding.
- If there are no configured source entries for the reporting group, the IGMP process adds to the interface that the report is received on to an (*,G) entry in the MRIB. This is the typical IGMP functionality.
- The periodic group reports helps to keep the state of the translated (S,G) alive. If there are no incoming reports, all entries time out at the same time.
- If an IGMPv2 leave message is received for the group and a corresponding translated entry exist, all entries expire at the same time unless an overriding report is received.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`Device# configure terminal`
	`Device(config)#`
		Enters configuration mode.

Step 2	**ip igmp ssm-translate** *group-prefix source-addr* **Example:** `Device(config)# ip igmp ssm-translate 232.0.0.0/8 10.1.1.1`	Configures the translation of IGMPv1 or IGMPv2 membership reports by the IGMP process to create the (S,G) state as if the router had received an IGMPv3 membership report.
Step 3	**show running-configuration igmp** **Example:** `Device(config)# show running-configuration igmp`	(Optional) shows the running-configuration information, including *ssm-translate* command lines.
Step 4	**show ip igmp groups** **Example:** `Device(config)# show ip igmp groups`	(Optional) Displays the IGMP attached group membership for a group or interface, the default VRF, a selected VRF, or all VRFs.
Step 5	**show ip mroute** **Example:**	(Optional) Shows IP multicast routing table for default VRF.

	Command or Action	Purpose
	`Device(config)# show ip mroute`	
Step 6	**copy running-config startup-config** **Example:** `Device(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring the Enforce Router Alert Option Check

You can configure the enforce router alert option check for IGMPv2 and IGMPv3 packets.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip igmp enforce-router-alert** **Example:** `switch(config)# ip igmp enforce-router-alert`	Enables the enforce router alert option check for IGMPv2 and IGMPv3 packets. By default, the enforce router alert option check is enabled.
Step 3	**no ip igmp enforce-router-alert** **Example:** `switch(config)# no ip igmp enforce-router-alert`	Disables the enforce router alert option check for IGMPv2 and IGMPv3 packets. By default, the enforce router alert option check is enabled.
Step 4	**show running-configuration igmp** **Example:** `switch(config)# show running-configuration igmp`	(Optional) Displays the running-configuration information, including the *enforce-router-alert* command line.
Step 5	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Restarting the IGMP Process

You can restart the IGMP process and optionally flush all routes.

Procedure

	Command or Action	Purpose
Step 1	**restart igmp** **Example:** `switch# restart igmp`	Restarts the IGMP process.

Step 2	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 3	**ip igmp flush-routes** **Example:** `switch(config)# ip igmp flush-routes`	Removes routes when the IGMP process is restarted. By default, routes are not flushed.
Step 4	**show running-configuration igmp** **Example:** `switch(config)# show running-configuration igmp`	(Optional) Displays the running-configuration information, including the *flush-routes* command lines.
Step 5	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Verifying the IGMP Configuration

To display the IGMP configuration information, perform one of the following tasks:

Command	Description
show ip igmp interface [*interface*] [**vrf** *vrf-name* | **all**] [**brief**]	Displays IGMP information about all interfaces or a selected interface, the default VRF, a selected VRF, or all VRFs. If IGMP is in vPC mode. Use this command to display vPC statistics.
show ip igmp groups [{**source** [*group*]}] | {**group** [*source*]}] [**interface**] [**summary**] [**vrf** *vrf-name* | **all**]	Displays the IGMP attached group membership for a group or interface, the default VRF, a selected VRF, or all VRFs.
show ip igmp route [{**source** [*group*]}] | {**group** [*source*]}] [**interface**] [**summary**] [**vrf** *vrf-name* | **all**]	Displays the IGMP attached group membership for a group or interface, the default VRF, a selected VRF, or all VRFs.
show ip igmp local- groups	Displays the IGMP local group membership.
show running-configuration igmp	Displays the IGMP running-configuration information.

Command	Description
show startup-configuration igmp	Displays the IGMP startup-configuration information.

For detailed information about the fields in the output from these commands, see the *Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*.

Configuration Examples for IGMP

The following example shows how to configure the IGMP parameters:

```
config t
  ip igmp ssm-translate 232.0.0.0/8 10.1.1.1
  interface ethernet 2/1
    ip igmp version 3
    ip igmp join-group 230.0.0.0
    ip igmp startup-query-interval 25
    ip igmp startup-query-count 3
    ip igmp robustness-variable 3
    ip igmp querier-timeout 300
    ip igmp query-timeout 300
    ip igmp query-max-response-time 15
    ip igmp query-interval 100
    ip igmp last-member-query-response-time 3
    ip igmp last-member-query-count 3
    ip igmp group-timeout 300
    ip igmp report-link-local-groups
    ip igmp report-policy my_report_policy
    ip igmp access-group my_access_policy
```

Feature History for IGMP

This table lists the release history for this feature.

Table 6: Feature History for IGMP

Feature Name	Releases	Feature Information
IGMP vPC Incremental Sync	6.2(2)	The **show ip igmp internal vpc** command was introduced.
ip igmp groups and **ip igmp route** commands.	6.1(1)	Commands updated with summary parameter. • **ip igmp groups** • **ip igmp route**

Feature Name	Releases	Feature Information
vPC	4.1(3)	Displays vPC statistics with the **show ip igmp interface**command. The following section provides information about this feature: • *Verifying the IGMP Configuration*.
Immediate Leave	4.1(3)	Minimizes the leave latency of IGMPv2 or MLDv1 group memberships on a given IGMP or MLD interface because the device does not send group-specific queries. For more information, see *Configuring IGMP Interface Parameters*.

CHAPTER 4

Configuring MLD

This chapter describes how to configure the Multicast Listener Discovery (MLD) on Cisco NX-OS devices for IPv6 networks.

Information About MLD

MLD is an IPv6 protocol that a host uses to request multicast data for a particular group. Using the information obtained through MLD, the software maintains a list of multicast group or channel memberships on a per-interface basis. The devices that receive MLD packets send the multicast data that they receive for requested groups or channels out the network segment of the known receivers.

MLDv1 is derived from IGMPv2, and MLDv2 is derived from IGMPv3. IGMP uses IP Protocol 2 message types, while MLD uses IP Protocol 58 message types, which is a subset of the ICMPv6 messages.

The MLD process is started automatically on the device. You cannot enable MLD manually on an interface. MLD is automatically enabled when you perform one of the following configuration tasks on an interface:

- Enable PIM6
- Statically bind a local multicast group
- Enable link-local group reports

MLD Versions

The device supports MLDv1 and MLDv2. MLDv2 supports MLDv1 listener reports.

By default, the software enables MLDv2 when it starts the MLD process. You can enable MLDv1 on interfaces where you want only its capabilities.

MLDv2 includes the following key changes from MLDv1:

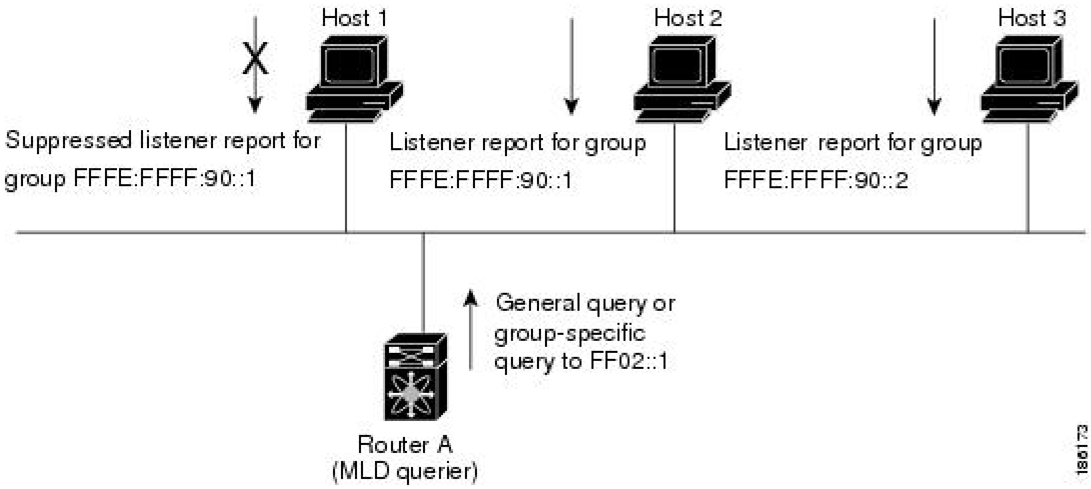
- Support for Source-Specific Multicast (SSM), which builds shortest path trees from each receiver to the source, through the following features:
    - Host messages that can specify both the group and the source.
    - The multicast state that is maintained for groups and sources, not just for groups as in MLDv1.
- Hosts no longer perform report suppression, which means that hosts always send MLD listener reports when an MLD query message is received.

For detailed information about MLDv1, see *RFC 2710.*For detailed information about MLDv2, see *RFC 3810.*

MLD Basics

The basic MLD process of a router that discovers multicast hosts is shown in the figure below. Hosts 1, 2, and 3 send unsolicited MLD listener report messages to initiate receiving multicast data for a group or channel.

Figure 12: MLD Query-Response Process


Host 1
Host 2
Host 3
Suppressed listener report for group FFFE:FFFF:90::1
Listener report for group FFFE:FFFF:90::1
Listener report for group FFFE:FFFF:90::2
General query or group-specific query to FF02::1
Router A (MLD querier)
186173

In this figure, router A, which is the MLD designated querier on the subnet, sends a general query message to the link-scope all-nodes multicast address FF02::1 periodically to discover what multicast groups hosts want to receive. The group-specific query is used to discover whether a specific group is requested by any hosts. You can configure the group membership timeout value that the router uses to determine that no members of a group or source exist on the subnet.

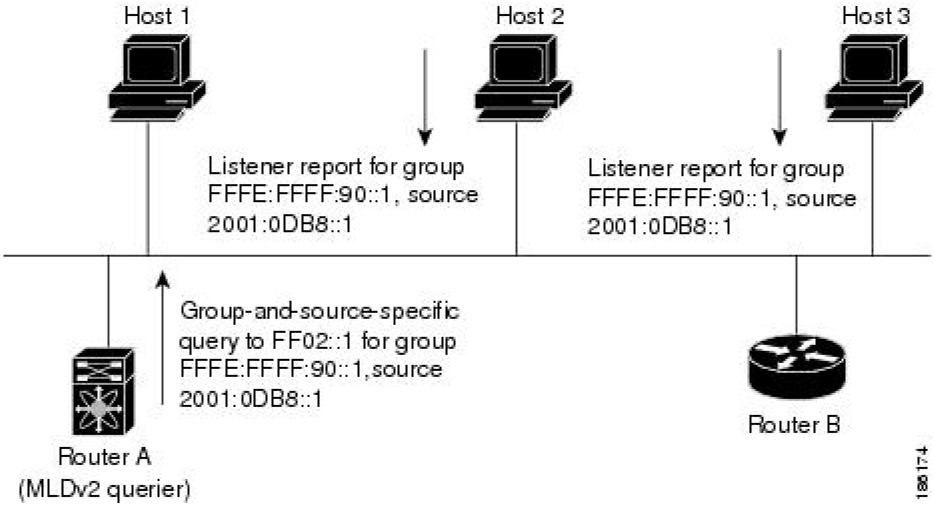
In this figure, host 1’s listener report is suppressed, and host 2 sends its listener report for group FFFE:FFFF:90::1 first. Host 1 receives the report from host 2. Because only one listener report per group needs to be sent to the router, other hosts suppress their reports to reduce network traffic. Each host waits for a random time interval to avoid sending reports at the same time. You can configure the query maximum response time parameter to control the interval in which hosts randomize their responses.

Note MLDv1 membership report suppression occurs only on hosts that are connected to the same port.

In this figure, router A sends the MLDv2 group-and-source-specific query to the LAN. Hosts 2 and 3 respond to the query with listener reports that indicate that they want to receive data from the advertised group and source. This MLDv2 feature supports SSM.

Note In MLDv2, all hosts respond to queries.

Figure 13: MLDv2 Group-and-Source-Specific Query


Host 1
Host 2
Host 3
Listener report for group FFFE:FFFF:90::1, source 2001:0DB8::1
Listener report for group FFFE:FFFF:90::1, source 2001:0DB8::1
Group-and-source-specific query to FF02::1 for group FFFE:FFFF:90::1,source 2001:0DB8::1
Router A (MLDv2 querier)
Router B
185174

The software elects a router as the MLD querier on a subnet if it has the lowest IP address. As long as a router continues to receive query messages from a router with a lower IP address, it remains a nonquerier and resets a timer that is based on its querier timeout value. If the querier timer of a router expires, it becomes the designated querier. If that router later receives a host query message from a router with a lower IP address, it drops its role as the designated querier and sets its querier timer again.

Messages sent by the designated querier have a time-to-live (TTL) value of 1, which means that the messages are not forwarded by the directly connected routers on the subnet, and you can configure the frequency and number of query messages sent specifically for MLD startup. You can configure a short query interval at startup so that the group state is established as quickly as possible. Although usually unnecessary, you can tune the query interval used after startup to a value that balances responsiveness to host group membership and the traffic created on the network.

Caution If you change the query interval, you can severely impact multicast forwarding in your network.

When a multicast host leaves a group, it should send a done message for MLDv1, or a listener report that excludes the group to the link-scope all-routers multicast address FF02::2. To check if this host is the last host to leave the group, the software sends an MLD query message and starts a timer that you can configure called the last member query response interval. If no reports are received before the timer expires, the software removes the group state. The router continues to send multicast traffic for a group until its state is removed.

You can configure a robustness value to compensate for the packet loss on a congested network. The robustness value is used by the MLD software to determine the number of times to send messages.

Link local addresses in the range FF02::0/16 have link scope, as defined by the Internet Assigned Numbers Authority (IANA). Network protocols on a local network segment use these addresses; routers do not forward these addresses because they have a TTL of 1. By default, the MLD process sends listener reports only for nonlink local addresses, but you can configure the software to send reports for link local addresses.

Virtualization Support

A virtual device context (VDC) is a logical representation of a set of system resources. Within each VDC, you can define multiple virtual routing and forwarding (VRF) instances. One MLD process can run per VDC. The MLD process supports all VRFs in that VDC.

For information about configuring VDCs, see the*Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*.

For information about configuring VRFs, see *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

Prerequisites for MLD

MLD has the following prerequisites:

- You are logged onto the device.
- You are in the correct virtual device context (VDC). A VDC is a logical representation of a set of system resources. You can use the **switchto vdc** command with a VDC number.
- For global configuration commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.

Guidelines and Limitations for MLD

MLD has the following guidelines and limitations:

- On M1, M2 and M3 Series modules, you must disable IGMP optimized multicast flooding (OMF) on all VLANs that require IPv6 multicast packet forwarding. To disable OMF, use the **no ip igmp snooping optimise-multicast-flood** command in VLAN configuration mode.
- On F2 Series modules, you must disable IGMP optimized multicast flooding (OMF) on all VLANs that require IPv6 packet forwarding (unicast or multicast). IPv6 neighbor discovery only functions in a VLAN with the OMF feature disabled. To disable OMF, use the **no ip igmp snooping optimise-multicast-flood** command in VLAN configuration mode.

Note When OMF is disabled, unknown IPv4 multicast traffic and all IPv6 multicast traffic is flooded to all ports in the VLAN.

(Unknown multicast traffic refers to multicast packets that have an active source, but have no receivers in the ingress VLAN. Having no receivers means that there is no group forwarding entry in the hardware.)

Default Settings for MLD

Table 7: Default MLD Parameters

Parameters	Default
MLD version	2
Startup query interval	30 seconds
Startup query count	2
Robustness value	2
Querier timeout	255 seconds
Query timeout	255 seconds
Query max response time	10 seconds
Query interval	125 seconds
Last member query response interval	1 second
Last member query count	2
Group membership timeout	260 seconds
Report link local multicast groups	Disabled
Immediate leave	Disabled

Configuring MLD Parameters

You can configure the MLD global and interface parameters to affect the operation of the MLD process.

Note Before you can access the MLD commands, you must enable the MLD feature.

Note If you are familiar with the Cisco IOS CLI, be aware that the Cisco NX-OS commands for this feature might differ from the Cisco IOS commands that you would use.

Configuring MLD Interface Parameters

Table 8: MLD Interface Parameters

Parameter	Description
MLD version	MLD version that is enabled on the interface. MLDv2 supports MLDv1. The MLD version can be 1 or 2. The default is 2.
Static multicast groups	Multicast groups that are statically bound to the interface. You can configure the groups to join the interface with the (*, G) state or specify a source IP to join with the (S, G) state. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. **Note** Although you can configure the (S, G) state, the source tree is built only if you enable MLDv2. You can configure a multicast group on all the multicast-capable routers on the network so that pinging the group causes all the routers to respond.
Static multicast groups on OIF	Multicast groups that are statically bound to the output interface. You can configure the groups to join the output interface with the (*, G) state or specify a source IP to join with the (S, G) state. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. Although you can configure the (S, G) state, the source tree is built only if you enable MLDv2.
Startup query interval	Startup query interval. By default, this interval is shorter than the query interval so that the software can establish the group state as quickly as possible. Values range from 1 to 18,000 seconds. The default is 30 seconds.
Startup query count	Number of queries sent at startup that are separated by the startup query interval. Values range from 1 to 10. The default is 2.

Parameter	Description
Robustness value	Robustness variable that you can tune to reflect expected packet loss on a congested network. You can increase the robustness variable to increase the number of times that packets are resent. Values range from 1 to 7. The default is 2.
Querier timeout	Number of seconds that the software waits after the previous querier has stopped querying and before it takes over as the querier. Values range from 1 to 65,535 seconds. The default is 255 seconds.
Query max response time	Maximum response time advertised in MLD queries. You can tune the burstiness of MLD messages on the network by setting a larger value so that host responses are spread out over a longer time. This value must be less than the query interval. Values range from 1 to 25 seconds. The default is 10 seconds.
Query interval	Frequency at which the software sends MLD host query messages. You can tune the number of MLD messages on the network by setting a larger value so that the software sends MLD queries less often. Values range from 1 to 18,000 seconds. The default is 125 seconds.
Last member query response interval	Query interval for response to an MLD query that the software sends after receiving a host leave message from the last known active host on the subnet. If no reports are received in the interval, the group state is deleted. You can use this value to tune how quickly the software stops transmitting on the subnet. The software can detect the loss of the last member of a group or source more quickly when the values are smaller. Values range from 1 to 25 seconds. The default is 1 second.
Last member query count	Number of times that the software sends an MLD query, separated by the last member query response interval, in response to a host leave message from the last known active host on the subnet. Values range from 1 to 5. The default is 2. **Caution** Setting this value to 1 means that a missed packet in either direction causes the software to remove the multicast state from the queried group or channel. The software can wait until the next query interval before the group is added again.

Parameter	Description
Group membership timeout	Group membership interval that must pass before the router decides that no members of a group or source exist on the network. Values range from 3 to 65,535 seconds. The default is 260 seconds.
Report link local multicast groups	Option that enables sending reports for groups in FF02::0/16. Link local addresses are used only by protocols on the local network. Reports are always sent for nonlink local groups. The default is disabled.
Report policy	Access policy for MLD reports that is based on a route-map policy.
Access groups	Option that configures a route-map policy to control the multicast groups that hosts on the subnet serviced by an interface can join. **Note** Only the **match ip multicast group** command is supported in this route map policy. The **match ip address** command for matching an ACL is not supported.
Immediate leave	Option that minimizes the leave latency of MLDv1 group memberships on a given MLD interface because the device does not send group-specific queries. When immediate leave is enabled, the device will remove the group entry from the multicast routing table immediately upon receiving a leave message for the group. The default is disabled. **Note** Use this command only when there is one receiver behind the interface for a given group.

[2] To configure route-map policies, see the Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters configuration mode.

Step 2	**interface** *interface*
	Example:
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enters interface mode on the interface type and number, such as *ethernet*

	Command or Action	Purpose
Step 3		The following commands can be used to configure the MLD interface parameters.

Option	Description
ipv6 mld version *value* Example `switch(config-if)# ipv6 mld version 2`	Sets the MLD version to the value specified. Values can be 1 or 2. The default is 2. The *no* form of the command sets the version to 2.

ipv6 mld join-group {**group** [**source** *source*] | **route-map** *policy-name*}
Example
`switch(config-if)# ipv6 mld join-group FFFE::1`
	Statically binds a multicast group to the interface. If you specify only the group address, the (*, G) state is created. If you specify the source address, the (S, G) state is created. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command.
	Note A source tree is built for the (S, G) state only if you enable MLDv2.
	Caution The device CPU must handle the traffic generated by using this command.

ipv6 mld static-oif {**group** [**source** *source*] | **route-map** *policy-name*}
Example
`switch(config-if)# ipv6 mld static-oif FFFE::1`
	Statically binds a multicast group to the outgoing interface, which is handled by the device hardware. If you specify only the group address, the (*, G) state is created. If you specify the source address, the (S, G) state is

	Command or Action	Purpose
	Option	**Description**
		created. You can specify a route-map policy name that lists the group prefixes, group ranges, and source prefixes to use with the **match ip multicast** command. **Note** A source tree is built for the (S, G) state only if you enable MLDv2.
	ipv6 mld startup-query-interval *seconds* Example `switch(config-if)# ipv6 mld startup-query-interval 25`	Sets the query interval used when the software starts up. Values can range from 1 to 18,000 seconds. The default is 31 seconds.
	ipv6 mld startup-query-count *count* Example `switch(config-if)# ipv6 mld startup-query-count 3`	Sets the query count used when the software starts up. Values can range from 1 to 10. The default is 2.
	ipv6 mld robustness-variable *value* Example `switch(config-if)# ipv6 mld robustness-variable 3`	Sets the robustness variable. You can use a larger value for a lossy network. Values can range from 1 to 7. The default is 2.
	ipv6 mld querier-timeout *seconds* Example `switch(config-if)# ipv6 mld querier-timeout 300`	Sets the querier timeout that the software uses when deciding to take over as the querier. Values can range from 1 to 65,535 seconds. The default is 255 seconds.
	ipv6 mld query-timeout *seconds* Example `switch(config-if)# ipv6 mld query-timeout 300`	Sets the query timeout that the software uses when deciding to take over as the querier. Values can range from 1 to

Command or Action	Purpose
Option	**Description**
	65,535 seconds. The default is 255 seconds. **Note** This command has the same functionality as the **ipv6 mld querier-timeout** command.
ipv6 mld query-max-response-time *seconds* Example `switch(config-if)# ipv6 mld query-max-response-time 15`	Sets the response time advertised in MLD queries. Values can range from 1 to 25 seconds. The default is 10 seconds.
ipv6 mld query-interval *interval* Example `switch(config-if)# ipv6 mld query-interval 100`	Sets the frequency at which the software sends MLD host query messages. Values can range from 1 to 18,000 seconds. The default is 125 seconds.
ipv6 mld last-member-query-response-time *seconds* Example `switch(config-if)# ipv6 mld last-member-query-response-time 3`	Sets the query interval waited after sending membership reports before the software deletes the group state. Values can range from 1 to 25 seconds. The default is 1 second.
ipv6 mld last-member-query-count *count* Example `switch(config-if)# ipv6 mld last-member-query-count 3`	Sets the number of times that the software sends an MLD query in response to a host leave message. Values can range from 1 to 5. The default is 2.
ipv6 mld group-timeout seconds Example `switch(config-if)# ipv6 mld group-timeout 300`	Sets the group membership timeout for MLDv2. Values can range from 3 to 65,535 seconds. The default is 260 seconds.
ipv6 mld report-link-local-groups	Enables sending reports for groups in

	Command or Action	Purpose

Option	Description
Example `switch(config-if)# ipv6 mld report-link-local-groups`	224.0.0.0/24. Reports are always sent for nonlink local groups. By default, reports are not sent for link local groups.

`ipv6 mld report-policy` *`policy`*
Example
`switch(config-if)# ipv6 mld report-policy my_report_policy`
	Configures an access policy for MLD reports that is based on a route-map policy.

`ipv6 mld access-group` *`policy`*
Example
`switch(config-if)# ipv6 mld access-group my_access_policy`
	Configures a route-map policy to control the multicast groups that hosts on the subnet serviced by an interface can join.
	Note Only the **match ip multicast group** command is supported in this route map policy. The **match ip address** command for matching an ACL is not supported.

`ipv6 mld immediate-leave`
Example
`switch(config-if)# ipv6 mld immediate-leave`
	Enables the device to remove the group entry from the multicast routing table immediately upon receiving a leave message for the group. Use this command to mnimize the leave latency of MLDv1 group memberships on a given MLD interface because the device does not send group-specific queries. The default is disabled.

	Command or Action	Purpose
	Option	**Description**
		Note Use this command only when there is one receiver behind the interface for a given group.
Step 4	**show ipv6 mld interface** [*interface*] **[vrf** *vrf-name* | **all**] [**brief**] **Example:** `switch(config)# show ipv6 mld interface`	(Optional) Displays MLD information about the interface.
Step 5	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring an MLD SSM Translation

You can configure an SSM translation to provide SSM support when the router receives MLDv1 listener reports. Only MLDv2 provides the capability to specify group and source addresses in listener reports. By default, the group prefix range is FF3x/96. To modify the PIM SSM range, see *Configuring SSM*.

Table 9: Example SSM Translations

Group Prefix	Source Address
FF30::0/16	2001:0DB8:0:ABCD::1
FF30::0/16	2001:0DB8:0:ABCD::2
FF30:30::0/24	2001:0DB8:0:ABCD::3
FF32:40::0/24	2001:0DB8:0:ABCD::4

This table shows the resulting M6RIB routes that the MLD process creates when it applies an SSM translation to the MLD v1 listener report. If more than one translation applies, the router creates the (S, G) state for each translation.

Table 10: Example Result of Applying SSM Translations

MLDv1 Listener Report	Resulting M6RIB Route
FF32:40::40	(2001:0DB8:0:ABCD::4, FF32:40::40)

MLDv1 Listener Report	Resulting M6RIB Route
FF30:10::10	(2001:0DB8:0:ABCD::1, FF30:10::10) (2001:0DB8:0:ABCD::2, FF30:10::10)

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ipv6** [icmp] **mld ssm-translate** *group-prefix source-addr* **Example:** `switch(config)# ipv6 mld ssm-translate FF30::0/16 2001:0DB8:0:ABCD::1`	Configures the translation of MLDv1 listener reports by the MLD process to create the (S,G) state as if the router had received an MLDv2 listener report.
Step 3	**show running-configuration ssm-translate** **Example:** `switch(config)# show running-configuration ssm-translate`	(Optional) Displays *ssm-translate* configuration lines in the running configuration.
Step 4	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Verifying the MLD Configuration

To display the MLD configuration information, perform one of the following tasks:

Command	Description
show ipv6 mld interface[`interface`] [**vrf** `vrf-name` | **all**] [brief]	Displays MLD information about all interfaces or a selected interface, the default VRF, a selected VRF, or all VRFs.
show ipv6 mld groups [`group` | `interface`] [**vrf** `vrf-name` | **all**]	Displays the MLD attached group membership for a group or interface, the default VRF, a selected VRF, or all VRFs.

show ipv6 mld route [**group** | **interface**] [**vrf** **vrf-name** | **all**]	Displays the MLD attached group membership for a group or interface, the default VRF, a selected VRF, or all VRFs.
show ipv6 mld local-groups-	Displays the MLD local group membership.

For detailed information about the fields in the output from these commands, see the *Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*.

Configuration Examples for MLD

The following example shows how to configure MLD:

```
config t
  ipv6 mld ssm-translate FF30::0/16 2001:0DB8:0:ABCD::1
  interface ethernet 2/1
    ipv6 mld version 2
    ipv6 mld join-group FFFE::1
    ipv6 mld startup-query-interval 25
    ipv6 mld startup-query-count 3
    ipv6 mld robustness-variable 3
    ipv6 mld querier-timeout 300
    ipv6 mld query-timeout 300
    ipv6 mld query-max-response-time 15
    ipv6 mld query-interval 100
    ipv6 mld last-member-query-response-time 3
    ipv6 mld last-member-query-count 3
    ipv6 mld group-timeout 300
    ipv6 mld report-link-local-groups
    ipv6 mld report-policy my_report_policy
    ipv6 mld access-group my_access_policy
```

Related Documents

Related Topic	Document Title
VDCs	*Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*
CLI commands	*Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*

Standards

Standards	Title
No new or modified standards are supported by this feature, and support for existing standards has not been modified by this feature.	—

Feature History for MLD

Table 11: Feature History for MLD

Feature Name	Releases	Feature Information
Immediate Leave	4.1(3)	Minimizes the leave latency of IGMPv2 or MLDv1 group memberships on a given IGMP or MLD interface because the device does not send group-specific queries. • *Configuring MLD Interface Parameters*

CHAPTER 5

Configuring PIM and PIM6

This chapter describes how to configure the Protocol Independent Multicast (PIM) and PIM6 features on Cisco NX-OS devices in your IPv4 and IPv6 networks.

Information About PIM and PIM6

Note Beginning with Cisco NX-OS Release 5.0(2a), Bidirectional Forwarding Detection (BFD) supports PIM. See the *Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide*.

PIM, which is used between multicast-capable routers, advertises group membership across a routing domain by constructing multicast distribution trees. PIM builds shared distribution trees on which packets from multiple sources are forwarded, as well as source distribution trees on which packets from a single source are forwarded. For more information about multicast, see *Information About Multicast*.

Cisco NX-OS supports PIM sparse mode for IPv4 networks (PIM) and for IPv6 networks (PIM6). In PIM sparse mode, multicast traffic is sent only to locations of the network that specifically request it. You can configure PIM and PIM6 to run simultaneously on a router. You can use PIM and PIM6 global parameters to configure RPs, message packet filtering, and statistics. You can use PIM and PIM6 interface parameters to enable multicast, identify PIM borders, set the PIM hello message interval, and set the designated router (DR) priority. For more information, see *Configuring PIM or PIM6 Sparse Mode*.

Note Cisco NX-OS does not support PIM dense mode.

In Cisco NX-OS, multicast is enabled only after you enable the PIM or PIM6 feature on each router and then enable PIM or PIM6 sparse mode on each interface that you want to participate in multicast. You can configure PIM for an IPv4 network and PIM6 for an IPv6 network. In an IPv4 network, if you have not already enabled IGMP on the router, PIM enables it automatically. In an IPv6 network, MLD is enabled by default. For information about configuring IGMP and MLD, see *Configuring IGMP* and *Configuring MLD*.

Note Beginning with Cisco NX-OS Release 5.2(1) for the Nexus 7000 Series devices, you can configure PIMv4 to run over generic routing encapsulation (GRE) tunnels including outgoing interfaces (OIFs).

You use the PIM and PIM6 global configuration parameters to configure the range of multicast group addresses to be handled by each of the three distribution modes:

- Any Source Multicast (ASM) provides discovery of multicast sources. It builds a shared tree between sources and receivers of a multicast group and supports switching over to a source tree when a new receiver is added to a group. ASM mode requires that you configure an RP.
- Single Source Multicast (SSM) builds a source tree originating at the designated router on the LAN segment that receives a request to join a multicast source. SSM mode does not require you to configure RPs. Source discovery must be accomplished through other means.
- Bidirectional shared trees (Bidir) build a shared tree between sources and receivers of a multicast group but do not support switching over to a source tree when a new receiver is added to a group. Bidir mode requires that you configure an RP. Bidir forwarding does not require source discovery because only the shared tree is used.

You can combine the three modes to cover different ranges of group addresses. For more information, see *Configuring PIM and PIM6*.

For more information about PIM sparse mode and shared distribution trees used by ASM and Bidir modes, see *RFC 4601*.

For more information about PIM SSM mode, see *RFC 3569*.

For more information about PIM Bidir mode, see *draft-ietf-pim-bidir-09.txt*.

Hello Messages

The PIM process begins when the router establishes PIM neighbor adjacencies by sending PIM hello messages to the multicast address 224.0.0.13. Hello messages are sent periodically at the interval of 30 seconds. When all neighbors have replied, the PIM software chooses the router with the highest priority in each LAN segment as the designated router (DR). The DR priority is based on a DR priority value in the PIM hello message. If the DR priority value is not supplied by all routers, or the priorities match, the highest IP address is used to elect the DR.

The hello message also contains a hold-time value, which is typically 3.5 times the hello interval. If this hold time expires without a subsequent hello message from its neighbor, the device detects a PIM failure on that link.

For added security, you can configure an MD5 hash value that the PIM software uses to authenticate PIM hello messages with PIM neighbors.

For information about configuring hello message authentication, see *Configuring PIM or PIM6 Sparse Mode*.

Join-Prune Messages

When the DR receives an IGMP membership report message from a receiver for a new group or source, the DR creates a tree to connect the receiver to the source by sending a PIM join message out the interface toward the rendezvous point (ASM or Bidir mode) or source (SSM mode).The rendezvous point (RP) is the root of a shared tree, which is used by all sources and hosts in the PIM domain in the ASM or the Bidir mode. SSM does not use an RP but builds a shortest path tree (SPT) that is the lowest cost path between the source and the receiver.

When the DR determines that the last host has left a group or source, it sends a PIM prune message to remove the path from the distribution tree.

The routers forward the join or prune action hop by hop up the multicast distribution tree to create (join) or tear down (prune) the path.

Note In this publication, the terms “PIM join message” and “PIM prune message” are used to simplify the action taken when referring to the PIM join-prune message with only a join or prune action.

Join-prune messages are sent as quickly as possible by the software. You can filter the join-prune messages by defining a routing policy. For information about configuring the join-prune message policy, see *Configuring PIM or PIM6 Sparse Mode*.

State Refreshes

PIM requires that multicast entries are refreshed within a 3.5-minute timeout interval. The state refresh ensures that traffic is delivered only to active listeners, and it keeps routers from using unnecessary resources.

To maintain the PIM state, the last-hop DR sends join-prune messages once per minute. State creation applies to both (*, G) and (S, G) states as follows:

- (*, G) state creation example—An IGMP (*, G) report triggers the DR to send a (*, G) PIM join message toward the RP.
- (S, G) state creation example—An IGMP (S, G) report triggers the DR to send an (S, G) PIM join message toward the source.

If the state is not refreshed, the PIM software tears down the distribution tree by removing the forwarding paths in the multicast outgoing interface list of the upstream routers.

Rendezvous Points

A rendezvous point (RP) is a router that you select in a multicast network domain that acts as a shared root for a multicast shared tree. You can configure as many RPs as you like, and you can configure them to cover different group ranges.

Static RP

You can statically configure an RP for a multicast group range. You must configure the address of the RP on every router in the domain.

You can define static RPs for the following reasons:

- To configure routers with the Anycast-RP address
- To manually configure an RP on a device

For information about configuring static RPs, see *Configuring Static RPs*.

BSRs

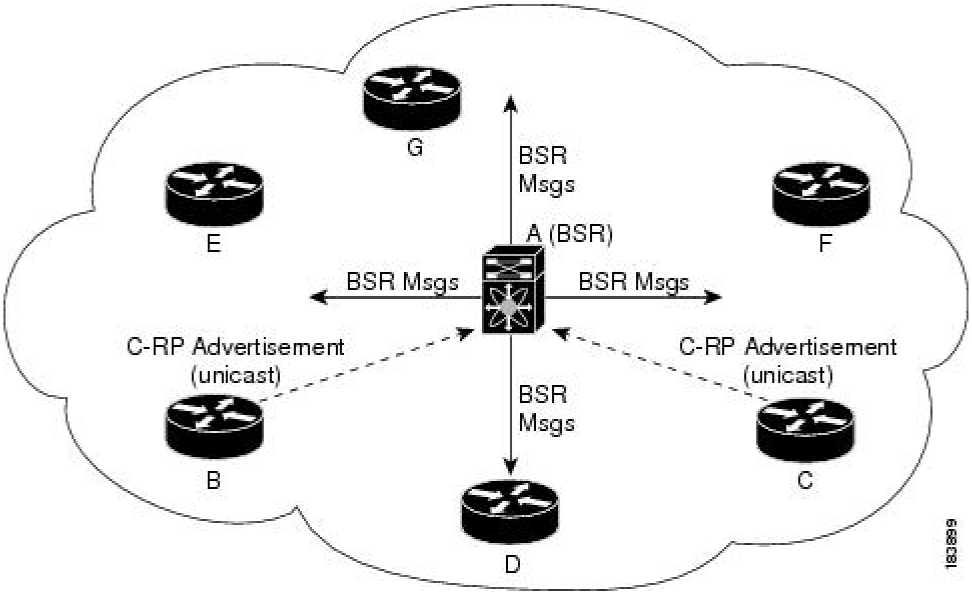
The bootstrap router (BSR) ensures that all routers in the PIM domain have the same RP cache as the BSR. You can configure the BSR to help you select an RP set from BSR candidate RPs. The function of the BSR is to broadcast the RP set to all routers in the domain. You select one or more candidate BSRs to manage the RPs in the domain. Only one candidate BSR is elected as the BSR for the domain.

Caution Do not configure both Auto-RP and BSR protocols in the same network.

This figure shows the BSR mechanism. Router A, the software-elected BSR, sends BSR messages out all enabled interfaces (shown by the solid lines in the figure). The messages, which contain the RP set, are flooded hop by hop to all routers in the network. Routers B and C are candidate RPs that send their candidate-RP advertisements directly to the elected BSR (shown by the dashed lines in the figure).

The elected BSR receives candidate-RP messages from all the candidate RPs in the domain. The bootstrap message sent by the BSR includes information about all of the candidate RPs. Each router uses a common algorithm to select the same RP address for a given multicast group.

Figure 14: BSR Mechanism



In the RP selection process, the RP address with the best priority is determined by the software. If the priorities match for two or more RP addresses, the software may use the RP hash in the selection process. Only one RP address is assigned to a group.

By default, routers are not enabled to listen or forward BSR messages. You must enable the BSR listening and forwarding feature so that the BSR mechanism can dynamically inform all routers in the PIM domain of the RP set assigned to multicast group ranges.

For more information about bootstrap routers, see *RFC 5059*.

Note The BSR mechanism is a nonproprietary method of defining RPs that can be used with third-party routers.

For information about configuring BSRs and candidate RPs, see *Configuring BSRs*.

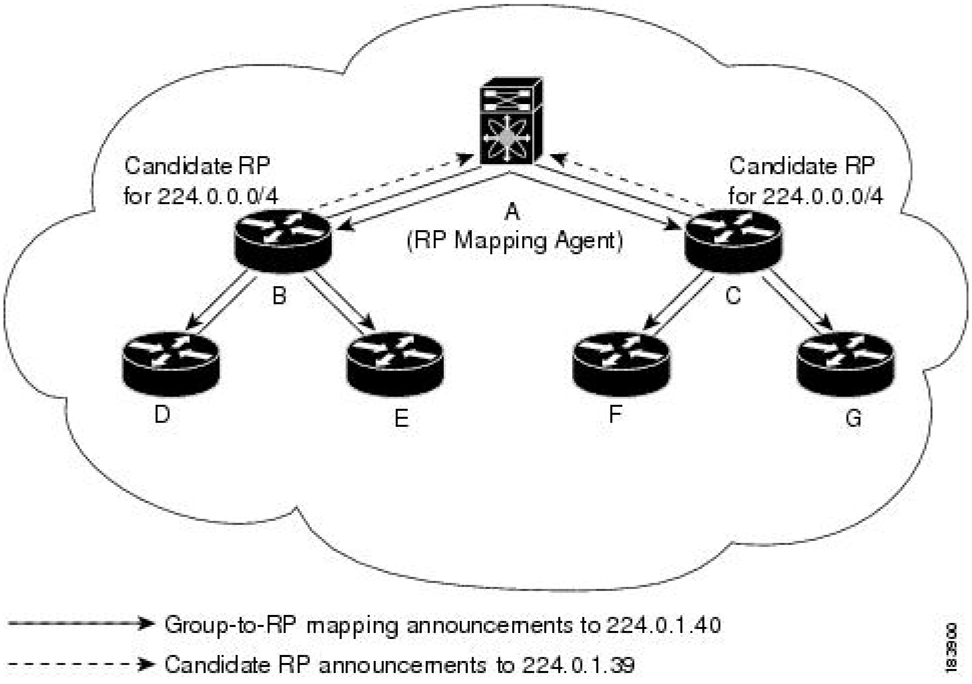
Auto-RP

Auto-RP is a Cisco protocol that was prior to the Internet standard bootstrap router mechanism. You configure Auto-RP by selecting candidate mapping agents and RPs. Candidate RPs send their supported group range in RP-Announce messages to the Cisco RP-Announce multicast group 224.0.1.39. An Auto-RP mapping agent listens for RP-Announce messages from candidate RPs and forms a Group-to-RP mapping table. The mapping agent multicasts the Group-to-RP mapping table in RP-Discovery messages to the Cisco RP-Discovery multicast group 224.0.1.40.

Caution Do not configure both Auto-RP and BSR protocols in the same network.

This figure shows the Auto-RP mechanism. Periodically, the RP mapping agent multicasts the RP information that it receives to the Cisco-RP-Discovery group 224.0.1.40 (shown by the solid lines in the figure).

Figure 15: Auto-RP Mechanism


Candidate RP for 224.0.0.0/4
A (RP Mapping Agent)
Candidate RP for 224.0.0.0/4
B
C
D
E
F
G
Group-to-RP mapping announcements to 224.0.1.40
Candidate RP announcements to 224.0.1.39
183900

By default, routers are not enabled to listen or forward Auto-RP messages. You must enable the Auto-RP listening and forwarding feature so that the Auto-RP mechanism can dynamically inform routers in the PIM domain of the group-to-RP mapping.

Note Auto-RP is not supported for PIM6.

For information about configuring Auto-RP, see *Configuring Auto-RP*.

Multiple RPs Configured in a PIM Domain

This section describes the election process rules when multiple RPs are configured in a PIM domain.

PIM BSR Bootstrap/Auto-RP Mapping-Agent Election Process

This section describes the BSR bootstrap Auto-RP mapping-agent election process.

Bootstrap Router (BSR) Election Process Details

- If the BSR priorities are different, the BSR with the highest priority (highest numerical value) is elected as the BSR router for the PIM domain (see configuration example 1).
  - Configuration example 1—Different BSR-candidate priorities: In this example, the system elects the device labeled N7K-1 as the BSR candidate for the PIM domain because it has the highest priority. The device labeled N7K-2 has the default priority of 64.

Configuration for N7K-1:

```
interface loopback0
  ip address 192.168.1.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0 priority 128

ip pim bsr forward listen
```

Configuration for N7K-2:

```
interface loopback0
  ip address 192.168.2.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0

ip pim bsr forward listen
```

Verification for N7K-1:

```
show ip pim rp
 PIM RP Status Information for VRF "default"
 BSR: 192.168.1.1*, next Bootstrap message in: 00:00:12,

      priority: 128, hash-length: 30
```

Verification for N7K-2:

```
show ip pim rp
 PIM RP Status Information for VRF "default"
 BSR: 192.168.1.1, uptime: 00:04:27, expires: 00:02:00,
      priority: 128, hash-length: 30
```

- If the BSR priorities are the same, the BSR with the highest BSR-candidate IP address is elected as the BSR router for the PIM domain (see configuration example 2).
  - Configuration example 2—Identical BSR-candidate priorities: In this example, the system elects the device labeled N7K-2 as the BSR for the PIM domain because it has the highest BSR-candidate IP address.

Configuration for N7K-1:

```
interface loopback0
  ip address 192.168.1.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0

ip pim bsr forward listen
```

Configuration for N7K-2:

```
interface loopback0
  ip address 192.168.2.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0

ip pim bsr forward listen
```

Verification for N7K-1:

```
show ip pim rp
PIM RP Status Information for VRF "default"
 BSR: 192.168.2.1, uptime: 01:45:20, expires: 00:01:54,
         priority: 64, hash-length: 30
```

Verification for N7K-2:

```
show ip pim rp
PIM RP Status Information for VRF "default"
 BSR: 192.168.2.1*, next Bootstrap message in: 00:00:30,
      priority: 64, hash-length: 30
```

Auto-RP Mapping Agent Election Process

- The router with the highest mapping-agent IP address is elected as the mapping agent for the PIM domain. You cannot configure the priority for the Auto-RP mapping agent (see configuration example):
    - Configuration example—Highest IP address: In this example, the system elects the device labeled N7K-2 as the mapping agent for the PIM domain because it has the highest mapping-agent IP address.

Configuration for N7K-1:

```
interface loopback0
  ip address 192.168.1.1/32
  ip pim sparse-mode

ip pim auto-rp mapping-agent loopback0

ip pim auto-rp forward listen
```

Configuration for N7K-2:

```
interface loopback0
  ip address 192.168.2.1/32
  ip pim sparse-mode

ip pim auto-rp mapping-agent loopback0

ip pim auto-rp forward listen
```

Verification for N7K-1:

```
show ip pim rp
PIM RP Status Information for VRF "default"
 BSR disabled
 Auto-RP RPA: 192.168.2.1, next Discovery message in: 00:00:52
```

Verification for N7K-2:

```
show ip pim rp
PIM RP Status Information for VRF "default"
 BSR disabled
 Auto-RP RPA: 192.168.2.1*, next Discovery message in: 00:00:47
```

PIM RP versus RP Election Process

This table shows the process that the system uses to select the RP for a multicast group if multiple RPs are configured in the network using BSR, Auto-RP, or static RP configurations.

BSR-RP vs. BSR-RP	BSR-RP vs. Static RP	Auto-RP vs. Auto- RP	Auto-RP vs. Static RP
1. Most specific RP group-list	1.Most specific RP group-list	1. Most specific RP group-list	1. Most specific RP group-list
2. Lowest RP priority	2. Highest RP IP address	2. Highest RP IP address	2. Highest RP IP address
3. Highest RP IP address	—	—	—

Note BSR-RP versus Auto-RP is not listed in this table because we recommend that you do not run both simultaneously in the same network.

PIM BSR RP-Candidate Versus BSR RP-Candidate Election Process

- The BSR RP-candidate with the most specific group list is elected as the RP for any multicast addresses specified in its configured group list. The most specific group list takes priority over the BSR RP-candidate priority and the highest BSR RP-candidate IP address (see configuration example 1).
    - Configuration example 1—Most specific group list: In this example, the system elects the device labeled N7K-1 as the RP for all multicast addresses specified in the 224.1.1.0/24 group-list. The system elects the device labeled N7K-2 for the multicast addresses within the less specific 224.0.0.0/4 group list.

Configuration for N7K-1:

```
interface loopback0
  ip address 192.168.1.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0
ip pim bsr rp-candidate loopback0 group-list 224.1.1.0/24
ip pim bsr forward listen
```

Configuration for N7K-2:

```
interface loopback0
  ip address 192.168.2.1/32
  ip pim sparse-mode
ip pim bsr bsr-candidate loopback0
ip pim bsr rp-candidate loopback0 group-list 224.0.0.0/4
ip pim bsr forward listen
```

Verification for N7K-1:

```
show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.1.1.0/24       ASM       192.168.1.1      -

show ip pim group 224.3.0.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.2.1      -
```

Verification for N7K-2:

```
show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.1.1.0/24       ASM       192.168.1.1      -

show ip pim group 224.3.0.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.2.1
```

-
- When multiple BSR RP-candidates advertise the same group list (for example, 224.0.0.0/4), the system elects the BSR RP-candidate with the highest priority (lowest numerical value) as the RP for any multicast address specified in its group list (see configuration example 2).
    - Configuration example 2—Identical group list with different RP priorities: In this example, the system elects the device labeled N7K-1 as the RP for all multicast addresses specified in the 224.0.0.0/4 group list because it has the lowest RP-candidate priority. The device labeled N7K-2 has a default priority of 192.

Configuration for N7K-1:

```
interface loopback0
  ip address 192.168.1.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0
ip pim bsr rp-candidate loopback0 group-list 224.0.0.0/4 priority 10
ip pim bsr forward listen
```

Configuration for N7K-2:

```
interface loopback0
  ip address 192.168.2.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0
ip pim bsr rp-candidate loopback0 group-list 224.0.0.0/4
ip pim bsr forward listen
```

Verification for N7K-1:

```
show ip pim rp
 PIM RP Status Information for VRF "default"
 BSR: 192.168.2.1, uptime: 00:09:14, expires: 00:01:37,
  priority: 64, hash-length: 30
 Auto-RP disabled
 BSR RP Candidate policy: None
 BSR RP policy: None
 Auto-RP Announce policy: None
 Auto-RP Discovery policy: None

 RP: 192.168.1.1*, (0), uptime: 00:08:15, expires: 00:01:57,
  priority: 10, RP-source: 192.168.2.1 (B), group ranges:
 224.0.0.0/4

 RP: 192.168.2.1, (0), uptime: 00:08:15, expires: 00:01:57,
  priority: 192, RP-source: 192.168.2.1 (B), group ranges:
 224.0.0.0/4

show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.1.1
```

Verification for N7K-2:

```
show ip pim rp
 PIM RP Status Information for VRF "default"
 BSR: 192.168.2.1*, next Bootstrap message in: 00:00:55,
  priority: 64, hash-length: 30
 Auto-RP disabled
 BSR RP Candidate policy: None
 BSR RP policy: None
 Auto-RP Announce policy: None
 Auto-RP Discovery policy: None

 RP: 192.168.1.1, (0), uptime: 00:11:34, expires: 00:02:26,
  priority: 10, RP-source: 192.168.1.1 (B), group ranges:
 224.0.0.0/4

 RP: 192.168.2.1*, (0), uptime: 00:12:21, expires: 00:02:22,
  priority: 192, RP-source: 192.168.2.1 (B), group ranges:
 224.0.0.0/4

show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.1.1      -
```

- When multiple BSR RP-candidates advertise the same group list (for example, 224.0.0.0/4) and are configured with the same BSR RP-candidate priority, the system elects the BSR RP-candidate with the highest IP address as the RP for any multicast address specified in its group list (see configuration example 3).
    - Configuration example 3—Identical group list with identical RP priorities: In this example, the system elects the device labeled N7K-2 as the RP for all multicast addresses specified in the 224.0.0.0/4 group list because it has the highest RP-candidate IP address.

Configuration for N7K-1:

```
interface loopback0
  ip address 192.168.1.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0
ip pim bsr rp-candidate loopback0 group-list 224.0.0.0/4
ip pim bsr forward listen
```

Configuration for N7K-2:

```
interface loopback0
  ip address 192.168.2.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0
ip pim bsr rp-candidate loopback0 group-list 224.0.0.0/4
ip pim bsr forward listen
```

Verification for N7K-1:

```
show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.2.1      -
```

Verification for N7K-2:

```
show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.2.1      -
```

PIM BSR RP-Candidate Versus Static RP Election Process

- The RP with the most specific group list is elected as the RP for any multicast addresses specified in its configured group list. The most specific group list takes priority over the highest RP IP address (see configuration example 1). (RP priorities are not applicable when comparing BSR RP-candidates to static RPs.)
  - Configuration example 1—Most specific group list: In this example, the system elects the device labeled N7K-1 as the BSR RP for all multicast addresses specified in the 224.1.1.0/24 group list. The system elects the device labeled N7K-2 as the RP for the multicast addresses within the less specific 224.0.0.0/4 group list because of the static RP statement.

Configuration for N7K-1:

```
interface loopback0
  ip address 192.168.1.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0
ip pim rp-address 192.168.2.1 group-list 224.0.0.0/4
ip pim bsr rp-candidate loopback0 group-list 224.1.1.0/24
ip pim forward listen
```

Configuration for N7K-2:

```
interface loopback0
  ip address 192.168.2.1/32
  ip pim sparse-mode

ip pim rp-address 192.168.2.1 group-list 224.0.0.0/4

ip pim bsr forward listen
```

Verification for N7K-1:

```
show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.1.1.0/24       ASM       192.168.1.1      -

show ip pim group 224.3.0.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.2.1      -
```

Verification for N7K-2:

```
show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.1.1.0/24       ASM       192.168.1.1      -

show ip pim group 224.3.0.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.2.1      -
```

- When a static RP and the BSR RP-candidate advertise the same group list (for example, 224.0.0.0/4), the system elects the system with the highest RP IP address as the RP for any multicast addresses specified in its group list (see configuration example 2).
  - Configuration example 2—Identical RP group list: In this example, the system elects the device labeled N7K-2 as the RP for all multicast addresses specified in the 224.0.0.0/4 group list because it has the highest RP IP address.

Configuration for N7K-1:

```
interface loopback0
  ip address 192.168.1.1/32
  ip pim sparse-mode

ip pim rp-address 192.168.1.1 group-list 224.0.0.0/4

ip pim bsr forward listen
```

Configuration for N7K-2:

```
interface loopback0
  ip address 192.168.2.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0
ip pim rp-address 192.168.1.1 group-list 224.0.0.0/4
ip pim bsr rp-candidate loopback0 group-list 224.0.0.0/4
ip pim bsr forward listen
```

Verification for N7K-1:

```
show ip pim group 224.1.1.0
PIM Group-Range Configuration for VRF "default"
Group-range        Mode      RP-address       Shared-tree-only range
224.0.0.0/4        ASM       192.168.2.1
```

Verification for N7K-2:

```
show ip pim group 224.1.1.0
PIM Group-Range Configuration for VRF "default"
Group-range        Mode      RP-address       Shared-tree-only range
224.0.0.0/4        ASM       192.168.2.1      -
```

- Because you cannot configure a static RP and its default value is 0, the RP priority has no impact. You can configure the BSR RP-candidate with a value between 0 and 255. The system elects the device with the most specific group list. If both devices have the same group list, the system elects the device with the highest RP IP address (see configuration example 3).
  - Configuration example 3—Identical group list and identical RP priorities: In this example, the system elects the device labeled N7K-2 as the RP for all multicast addresses specified in the 224.0.0.0/4 group list because it has the highest RP IP address. The system does not compare RP priorities between BSR RPs and static RPs.

Configuration for N7K-1:

```
interface loopback0
  ip address 192.168.1.1/32
  ip pim sparse-mode

ip pim bsr bsr-candidate loopback0
ip pim rp-address 192.168.2.1 group-list 224.0.0.0/4
ip pim bsr rp-candidate loopback0 group-list 224.0.0.0/4 priority 0

ip pim bsr forward listen
```

Configuration for N7K-2:

```
interface loopback0
  ip address 192.168.2.1/32
  ip pim sparse-mode

ip pim rp-address 192.168.2.1 group-list 224.0.0.0/4

ip pim bsr forward listen
```

Verification for N7K-1:

```
show ip pim rp
 PIM RP Status Information for VRF "default"
 BSR: 192.168.1.1*, next Bootstrap message in: 00:00:52,
  priority: 64, hash-length: 30
 Auto-RP disabled
 BSR RP Candidate policy: None
 BSR RP policy: None
 Auto-RP Announce policy: None
 Auto-RP Discovery policy: None

 RP: 192.168.1.1*, (0), uptime: 00:01:57, expires: 00:02:25,
 priority: 0, RP-source: 192.168.1.1 (B), group ranges:
  224.0.0.0/4
 RP: 192.168.2.1, (0), uptime: 02:16:09, expires: never,
 priority: 0, RP-source: (local), group ranges:
  224.0.0.0/4

show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.2.1      -
```

Verification for N7K-2:

```
show ip pim rp
 PIM RP Status Information for VRF "default"
 BSR: 192.168.1.1, uptime: 00:29:47, expires: 00:01:45,
  priority: 64, hash-length: 30
 Auto-RP disabled
 BSR RP Candidate policy: None
 BSR RP policy: None
 Auto-RP Announce policy: None
 Auto-RP Discovery policy: None

 RP: 192.168.1.1, (0), uptime: 00:06:59, expires: 00:02:05,
 priority: 0, RP-source: 192.168.1.1 (B), group ranges:
  224.0.0.0/4
 RP: 192.168.2.1*, (0), uptime: 00:13:15, expires: never,
 priority: 0, RP-source: (local), group ranges:
  224.0.0.0/4

show ip pim group 224.1.1.0
 PIM Group-Range Configuration for VRF "default"
 Group-range        Mode      RP-address       Shared-tree-only range

 224.0.0.0/4        ASM       192.168.2.1      -
```

PIM Auto-RP-Candidate Versus Auto-RP-Candidate Election Process

The auto-RP-candidate election is similar to the BSR RP-candidate election process, but it does not support priorities (see the *PIM BSR RP-Candidate vs. BSR RP-Candidate Election Process*). You cannot configure the priority for an auto-RP. The default value is 0.

PIM Auto-RP-Candidate Versus Static RP Election Process

The auto-RP-candidate versus static RP election uses the same rules as the election process for the BSR RP-candidate versus static RP See *PIM BSR RP-Candidate vs. Static RP Election Process*.

Anycast-RP

Anycast-RP has two implementations: one uses Multicast Source Discovery Protocol (MSDP) and the other is based on*RFC 4610, Anycast-RP Using Protocol Independent Multicast (PIM)*. This section describes how to configure PIM Anycast-RP.

You can use PIM Anycast-RP to assign a group of routers, called the Anycast-RP set, to a single RP address that is configured on multiple routers. The set of routers that you configure as Anycast-RPs is called the Anycast-RP set. This method is the only RP method that supports more than one RP per multicast group, which allows you to load balance across all RPs in the set. The Anycast RP supports all multicast groups.

PIM register messages are sent to the closest RP and PIM join-prune messages are sent in the direction of the closest RP as determined by the unicast routing protocols. If one of the RPs goes down, unicast routing ensures these message will be sent in the direction of the next-closest RP.

You must configue PIM on the loopback interface that is used for the PIM Anycast RP.

For more information about PIM Anycast-RP, see *RFC 4610*.

For information about configuring Anycast-RPs, see *Configuring a PIM Anycast-RP Set*.

PIM Register Messages

PIM register messages are unicast to the RP by designated routers (DRs) that are directly connected to multicast sources. The PIM register message has the following functions:

- To notify the RP that a source is actively sending to a multicast group.
- To deliver multicast packets sent by the source to the RP for delivery down the shared tree.

The DR continues to send PIM register messages to the RP until it receives a Register-Stop message from the RP. The RP sends a Register-Stop message in either of the following cases:

- The RP has no receivers for the multicast group being transmitted.
- The RP has joined the SPT to the source but has not started receiving traffic from the source.

You can use the **ip pim register-source** command to configure the IP source address of register messages when the IP source address of a register message is not a uniquely routed address to which the RP can send packets. This situation might occur if the source address is filtered so that the packets sent to it are not forwarded or if the source address is not unique to the network. In these cases, the replies sent from the RP to the source address will fail to reach the DR, resulting in Protocol Independent Multicast sparse mode (PIM-SM) protocol failures.

This example shows how to configure the IP source address of the register message to the loopback 3 interface of a DR:

```
ip pim register-source loopback 3
```

Note In Cisco NX-OS, PIM register messages are rate limited to avoid overwhelming the RP.

You can filter PIM register messages by defining a routing policy. For information about configuring the PIM register message policy, see the *Configuring Shared Trees Only for ASM*.

Designated Routers

In PIM ASM and SSM modes, the software chooses a designated router (DR) from the routers on each network segment. The DR is responsible for forwarding multicast data for specified groups and sources on that segment.

The DR for each LAN segment is determined as described in the *Hello Messages*.

In ASM mode, the DR is responsible for unicasting PIM register packets to the RP. When a DR receives an IGMP membership report from a directly connected receiver, the shortest path is formed to the RP, which may or may not go through the DR. The result is a shared tree that connects all sources transmitting on the same multicast group to all receivers of that group.

In SSM mode, the DR triggers (*, G) or (S, G) PIM join messages toward the RP or the source. The path from the receiver to the source is determined hop by hop. The source must be known to the receiver or the DR.

For information about configuring the DR priority, see the *Configuring PIM or PIM6 Sparse Mode*.

Designated Forwarders

In PIM Bidir mode, the software chooses a designated forwarder (DF) at RP discovery time from the routers on each network segment. The DF is responsible for forwarding multicast data for specified groups on that segment. The DF is elected based on the best metric from the network segment to the RP.

If the router receives a packet on the RPF interface toward the RP, the router forwards the packet out all interfaces in the OIF-list. If a router receives a packet on an interface on which the router is the elected DF for that LAN segment, the packet is forwarded out all interfaces in the OIF-list except the interface that it was received on and also out the RPF interface toward the RP.

Note Cisco NX-OS does not support PIM Bidir mode on F2 modules.

Note Cisco NX-OS puts the RPF interface into the OIF-list of the MRIB, but not in the OIF-list of the MFIB.

ASM Switchover from Shared Tree to Source Tree

Note Cisco NX-OS puts the RPF interface into the OIF-list of the MRIB, but not in the OIF-list of the MFIB.

In ASM mode, the DR that is connected to a receiver switches over from the shared tree to the shortest-path tree (SPT) to a source unless you configure the PIM parameter to use shared trees only. For information about configuring the use of shared trees only, see the *Configuring Shared Trees Only for ASM*.

During the switchover, messages on the SPT and shared tree may overlap. These messages are different. The shared tree messages are propagated upstream toward the RP, while SPT messages go toward the source.

For information about SPT switchovers, see the “Last-Hop Switchover” to the SPT section in *RFC 4601*.

ECMP Multicast Load Splitting Based on Source Group and Next-Hop Address Overview

The Advanced Multicast Multipath Support feature adds support for Equal Cost Multipath (ECMP) multicast load splitting based on source, group, and next-hop address. This feature enables multicast traffic from devices that send many streams to groups or that broadcast many channels, such as IPTV servers or MPEG video servers, to be more effectively load split across equal-cost paths.

Configuring ECMP multicast load splitting based on source, group, and next-hop address enables a more complex hash, the next-hop-based S-G-hash algorithm, which is based on source, group, and next-hop address. The next-hop-based S-G-hash algorithm is predictable because no randomization is used in calculating the hash value. Unlike the S-hash and basic S-G-hash algorithms, the hash mechanism used by the next-hop-based S-G-hash algorithm is not subject to polarization.

Note The next-hop-based S-G-hash algorithm in IPv4 multicast is the same algorithm used in IPv6 ECMP multicast load splitting, which, in turn, utilizes the same hash function used for PIM-SM bootstrap device (BSR).

The next-hop-based hash mechanism does not produce polarization and also maintains better RPF stability when paths fail. These benefits come at the cost that the source or RP IP addresses cannot be used to reliably predict and engineer the outcome of load splitting when the next-hop-based S-G-hash algorithm is used. Because many customer networks have implemented equal-cost multipath topologies, the manual engineering of load splitting, thus, is not a requirement in many cases. Rather, it is more of a requirement that the default behavior of IP multicast be similar to IP unicast; that is, it is expected that IP multicast use multiple equal-cost paths on a best-effort basis. Load splitting for IPv4 multicast, therefore, could not be enabled by default because of the anomaly of polarization.

Note Load splitting for CEF unicast also uses a method that does not exhibit polarization and likewise cannot be used to predict the results of load splitting or engineer the outcome of load splitting.

The next-hop-based hash function avoids polarization because it introduces the actual next-hop IP address of PIM neighbors into the calculation, so the hash results are different for each device, and in effect, there is no problem of polarization. In addition to avoiding polarization, this hash mechanism also increases stability of the RPF paths chosen in the face of path failures. Consider a device with four equal-cost paths and a large number of states that are load split across these paths. Suppose that one of these paths fails, leaving only three available paths. With the hash mechanism used by the polarizing hash mechanisms (the hash mechanism used by the S-hash and basic S-G-hash algorithms), the RPF paths of all states would likely reconverge and thus change between those three paths, especially those paths that were already using one of those three paths. These states, therefore, may unnecessarily change their RPF interface and next-hop neighbor. This problem exists simply because the chosen path is determined by taking the total number of paths available into consideration by the algorithm, so once a path changes, the RPF selection for all states is subject to change too. For the next-hop-based hash mechanism, only the states that were using the changed path for RPF would need to reconverge onto one of the three remaining paths. The states that were already using one of those paths would not change. If the fourth path came back up, the states that initially used it would immediately reconverge back to that path without affecting the other states.

Note The next-hop-based S-G-hash algorithm ignores bidir-PIM groups.

Administratively Scoped IP Multicast

The administratively scoped IP multicast method allows you to set boundaries on the delivery of multicast data. For more information, see *RFC 2365*.

You can configure an interface as a PIM boundary so that PIM messages are not sent out that interface. For information about configuring the domain border parameter, see the *Configuring PIM or PIM6 Sparse Mode*.

You can use the Auto-RP scope parameter to set a time-to-live (TTL) value. For more information, see the *Configuring Shared Trees Only for ASM*.

Bidirectional Forwarding Detection for PIM

Beginning with Cisco NX-OS Release 5.0(2a), Bidirectional Forwarding Detection (BFD) allows the system to rapidly detect failures in a network. See the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide, Release 6.x*, for more information about BFD.

In PIM, a link or neighbor group failure is detected when the hold-time, which is set as part of the hello interval, expires. However, BFD provides a more efficient method to detect a failure. This protocol establishes a session between the two endpoints over a link and uses the forwarding engine. When BFD is enabled, the PIM process attempts to add a BFD session as each neighbor is discovered. If a BFD session already exists, no duplicate is created but PIM receives a callback that contains the state of the BFD session. You can enable BFD for PIM per VRF or per interface.

PIM removes the BFD session when you disable BFD for that VRF or interface, the interface is no longer a PIM interface, or the neighboring BFD session goes down.

Virtualization Support

A virtual device context (VDC) is a logical representation of a set of system resources. Within each VDC, multiple virtual routing and forwarding (VRF) instances can be defined. For each VRF in a VDC in the system, independent multicast system resources are maintained, including the MRIB and M6RIB.

You can use the PIM and PIM6 **show** commands with a VRF argument to provide a context for the information displayed. The default VRF is used if no VRF argument is supplied.

For information about configuring VDCs, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*.

For information about configuring VRFs, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

Support for Graceful Restart PIM

The Support for Graceful Restart protocol independent multicast (PIM) feature is a multicast High Availability (HA) enhancement that improves the convergence of multicast-routes (mroutes) after a Route Processor (RP) switchover. In the event of an RP switchover, the support for Graceful Restart PIM feature utilizes the

Generation ID (GenID) value (defined in RFC 4601) as a mechanism to trigger adjacent PIM neighbors on an interface to send PIM join messages for all (*, G) and (S, G) states that use that interface as a reverse path forwarding (RPF) interface. This mechanism enables PIM neighbors to immediately reestablish those states on the newly active RP.

Prerequisites for Graceful Restart PIM

All Protocol Independent Multicast (PIM) neighbors must be compliant with RFC 4601 and be able to process Generation ID (GenID) differences in PIM hello messages.

Information About Graceful Restart PIM

Generation IDs

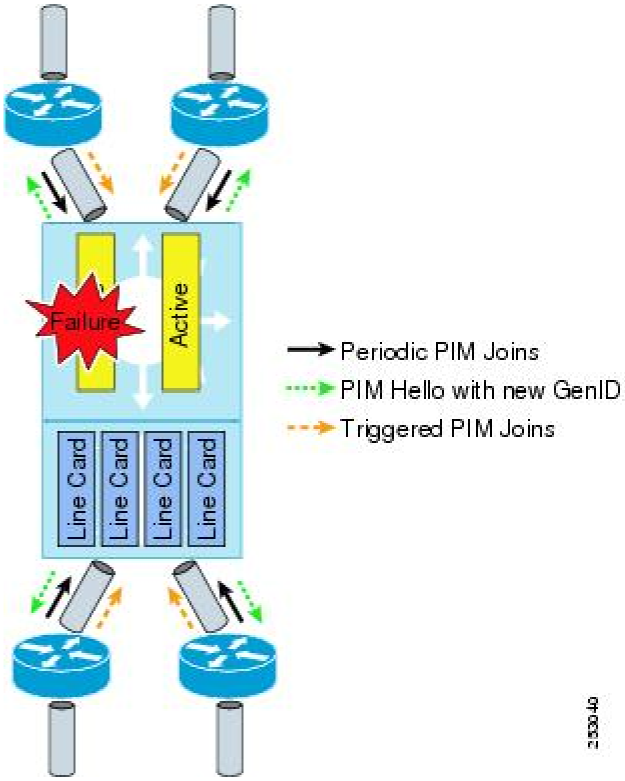
A Generation ID (GenID) is a randomly generated 32-bit value that is regenerated each time protocol independent multicast (PIM) forwarding is started or restarted on an interface. In order to process the GenID value in PIM hello messages, PIM neighbors must be running Cisco software with an implementation of PIM that is compliant with RFC 4601.

Note PIM neighbors that are not compliant with RFC 4601 and are unable to process GenID differences in PIM hello messages will ignore the GenIDs.

Graceful Restart PIM Functional Overview

The figure illustrates the operations that occur after a Route Processor (RP) switchover on devices that support the support for Graceful Restart protocol independent multicast (PIM) feature.

Figure 16: Operation of Graceful Restart PIM during an RP Switchover



The mechanics of the support for Graceful Restart PIM feature are as follows:

- In steady state, PIM neighbors exchange periodic PIM hello messages.
- An active RP receives PIM joins periodically to refresh multicast-route (mroute) states.
- When an active RP fails, the standby RP takes over to become the new active RP.
- The new active RP then modifies the Generation ID (GenID) value and sends the new GenID in PIM hello messages to adjacent PIM neighbors.
- Adjacent PIM neighbors that receive PIM hello messages on an interface with a new GenID send graceful restart PIM for all (*, G) and (S, G) mroutes that use that interfaces as an RPF interface.
- Those mroute states are then immediately reestablished on the newly active RP.

Graceful Restart PIM and Multicast Traffic Flow

Multicast traffic flow on PIM neighbors is not affected if the multicast traffic detects support for Graceful Restart PIM or PIM hello message from a node with the failing RP within the default PIM hello hold-time interval. Multicast traffic flow on a failing RP is not affected if it is Non-Stop Forwarding (NSF) capable.

Caution The default PIM hello hold-time interval is 3.5 times the PIM hello period. Multicast High Availability (HA) operations may not function as per design if you configure PIM hello interval with a value lower than the default value of 30 seconds.

Additional References for Graceful Restart PIM

RFCs

RFC	Title
RFC 4601	Protocol Independent Multicast - Sparse Mode (PIM-SM)

Technical Assistance

Description	Link
The Cisco Support website provides extensive online resources, including documentation and tools for troubleshooting and resolving technical issues with Cisco products and technologies. To receive security and technical information about your products, you can subscribe to various services, such as the Product Alert Tool (accessed from Field Notices), the Cisco Technical Services Newsletter, and Really Simple Syndication (RSS) Feeds. Access to most tools on the Cisco Support website requires a Cisco.com user ID and password.	http://www.cisco.com/support

High Availability

For information about high availability, see the *Cisco Nexus 7000 Series NX-OS High Availability and Redundancy Guide*.

Prerequisites for PIM and PIM6

PIM and PIM6 have the following prerequisites:

- You are logged onto the device.
- You are in the correct virtual device context (VDC). A VDC is a logical representation of a set of system resources. You can use the **switchto vdc** command with a VDC number.
- For global commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.

Guidelines and Limitations for PIM and PIM6

PIM and PIM6 have the following configuration guidelines and limitations:

- PIM must be configured on all Layer 3 interfaces between sources, receivers, and rendezvous points (RPs).

- Cisco NX-OS PIMv4 do not support route-map configuration with RP-Type. You can only configure Group Address, Source Address, and RP-address in the route-map.
- Tunnel interfaces do not support PIM until Cisco NX-OS Release 5.2(1). Beginning with Release 5.2(1), you can configure multicast on generic routing encapsulation (GRE) tunnel interfaces.
- The Cisco NX-OS software does not support multicast on a GRE tunnel interface that is in a different virtual routing and forwarding (VRF) instance than the VRF of the transport interface.
- Cisco NX-OS PIM and PIM6 do not interoperate with any version of PIM dense mode or PIM sparse mode version 1.
- Do not configure both Auto-RP and BSR protocols in the same network.
- Configure candidate RP intervals to a minimum of 15 seconds.
- If a device is configured with a BSR policy that should prevent it from being elected as the BSR, the device ignores the policy. This behavior results in the following undesirable conditions:
    - If a device receives a BSM that is permitted by the policy, the device, which incorrectly elected itself as the BSR, drops that BSM so that routers downstream fail to receive it. Downstream devices correctly filter the BSM from the incorrect BSR so that these devices do not receive RP information.
    - A BSM received by a BSR from a different device sends a new BSM but ensures that downstream devices do not receive the correct BSM.
- F2-Series modules do not support any form of IPv4 or IPv6 tunnels.
- Beginning with Release 5.x, using BFD for PIM to support fast failure detection is recommended.
- Default values for the PIM hello interval are recommended and should not be modified.

Note Aggressive PIM timers have been tested and can be supported in deployments where PIM timers must be modified. However this testing was limited and SSO/ISSU cannot be guaranteed in such a deployment. For more information, see the *Cisco Nexus 7000 Series NX-OS Verified Scalability Guide.*

- Cisco NX-OS PIM and PIM6 do not support Bidir PIM or SSM on vPCs.
- PIM adjacency with a vPC leg or with a router behind a vPC is not supported.

  A PIM adjacency between an Switched Virtual Interface (SVI) on a vPC VLAN (a VLAN that is carried on a vPC Peer-Link) and a downstream device is not supported; this configuration can result in dropped multicast packets. If a PIM neighbor relationship is required with a downstream device, a physical Layer 3 interface must be used on the Nexus switches instead of a vPC SVI.

  For SVIs on vPC Vlans, only one PIM adjacency is supported - which is with the vPC Peer Switch. PIM adjacencies over the VPC Peer-Link with devices other than the VPC Peer Switch for the vPC-SVI are NOT supported.
- Beginning with Release 7.1, PIM Bidir mode is not supported for VDCs that have the F2 Module. Bidir mode is supported on F2E or F2E with F3 modules on the same VDC.

- Use the **ip igmp static-oif** command on a Layer 3 interface of Cisco Nexus device to force the interface getting populated as an Outgoing Interface List (OIL). Do not use the **ip igmp join-group** command for this purpose.
- Multicast works on periodic joins/prune and depending on the topology and number of routers in the network, S,G state takes time to expire.
- The `sprase-mode` must be enabled by using the **ip pim sparse-mode** command on loopback interfaces that are configured as PIM rendezvous points.
- The interface that is used to configure a PIM RP (whether static, BSR or Auto-RP) must have **ip** [**v6**] **pim sparse-mode**.

Default Settings

Table 12: Default PIM and PIM6 Parameters

Parameters	Default
Use shared trees only	Disabled
Flush routes on restart	Disabled
Log Neighbor changes	Disabled
Auto-RP message action	Disabled
BSR message action	Disabled
SSM multicast group range or policy	232.0.0.0/8 for IPv4 and FF3x::/96 for IPv6
PIM sparse mode	Disabled
Designated router priority	0
Hello authentication mode	Disabled
Domain border	Disabled
RP address policy	No message filtering
PIM register message policy	No message filtering
BSR candidate RP policy	No message filtering
BSR policy	No message filtering
Auto-RP mapping agent policy	No message filtering
Auto-RP RP candidate policy	No message filtering
Join-prune policy	No message filtering
Neighbor adjacency policy	Become adjacent with all PIM neighbors

Parameters	Default
BFD	Disabled

Configuring PIM and PIM6

You can configure both PIM and PIM6 on the same router. You configure either PIM or PIM6 for each interface, depending on whether that interface is running IPv4 or IPv6.

Note Cisco NX-OS supports only PIM sparse mode version 2. In this publication, “PIM” refers to PIM sparse mode version 2.

You can configure separate ranges of addresses in the PIM or PIM6 domain using the multicast distribution modes described in the table below.

Multicast Distribution Mode	Requires RP Configuration	Description
ASM	Yes	Any source multicast
Bidir	Yes	Bidirectional shared trees
SSM	No	Single source multicast
RPF routes for multicast	No	RPF routes for multicast

PIM and PIM6 Configuration Tasks

The following steps configure PIM and PIM6.

1. From the multicast distribution modes, select the range of multicast groups that you want to configure in each mode.
2. From the multicast distribution modes, select the range of multicast groups that you want to configure in each mode.
3. Enable the PIM and PIM6 features.
4. Follow the configuration steps for the multicast distribution modes that you selected in Step 1.
    - For ASM or Bidir mode, see the *Configuring ASM and Bidir*.
    - For SSM mode, see the *Configuring SSM*.
    - For RPF routes for multicast, see the *Configuring RPF Routes for Multicast*.
5. Configure message filtering.

Note The CLI commands used to configure PIM or PIM6 differ as follows:

- Commands begin with **ip pim for PIM** and begin with **ipv6 pim for PIM6**
- Commands begin with **show ip pim** for PIM and begin with **show ipv6 pim** for PIM6.

Note If you are familiar with the Cisco IOS CLI, be aware that the Cisco NX-OS commands for this feature might differ from the Cisco IOS commands that you would use.

Enabling the PIM and PIM6 Features

Before you can access the PIM or PIM6 commands, you must enable the PIM or PIM6 feature.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**feature pim** **Example:** `switch(config)# feature pim`	Enables PIM. By default, PIM is disabled.
Step 3	**feature pim6** **Example:** `switch(config)# feature pim6`	Enables PIM6. By default, PIM6 is disabled.
Step 4	**show running-configuration pim** **Example:** `switch(config)# show running-configuration pim`	(Optional) Shows the running-configuration information for PIM, including the **feature** command.
Step 5	**show running-configuration pim6** **Example:** `switch(config)# show running-configuration pim6`	(Optional) Shows the running-configuration information for PIM6, including the **feature** command.
Step 6	**copy running-config startup-config** **Example:**	(Optional) Saves configuration changes.

	Command or Action	Purpose
	`switch(config)# copy running-config startup-config`	

Configuring PIM or PIM6 Sparse Mode Parameters

You configure PIM or PIM6 sparse mode on every device interface that you want to participate in a sparse mode domain. You can configure the sparse mode parameters described in the table below.

Table 13: PIM and PIM6 Sparse Mode Parameters

Parameter	Description
Global to the device	
Auto-RP message action	Enables listening and forwarding of Auto-RP messages. The default is disabled, which means that the router does not listen or forward Auto-RP messages unless it is configured as a candidate RP or mapping agent. **Note** PIM6 does not support the Auto-RP method.
BSR message action	Enables listening and forwarding of BSR messages. The default is disabled, which means that the router does not listen or forward BSR messages unless it is configured as a candidate RP or BSR candidate.
Bidir RP limit	Configures the number of Bidir RPs that you can configure for IPv4 and IPv6. The maximum number of Bidir RPs supported per VRF for PIM and PIM6 combined cannot exceed 8. Values range from 0 to 8. The default is 6 for IPv4 and 2 for IPv6.
Register rate limit	Configures the IPv4 or IPv6 register rate limit in packets per second. The range is from 1 to 65,535. The default is no limit.
Initial holddown period	Configures the IPv4 or IPv6 initial holddown period in seconds. This holddown period is the time it takes for the MRIB to come up initially. If you want faster convergence, enter a lower value. The range is from 90 to 210. Specify 0 to disable the holddown period. The default is 210.
Per device interface	
PIM sparse mode	Enables PIM or PIM6 on an interface.

Parameter	Description
Designated router priority	Sets the designated router (DR) priority that is advertised in PIM hello messages on this interface. On a multiaccess network with multiple PIM-enabled routers, the router with the highest DR priority is elected as the DR router. If the priorities match, the software elects the DR with the highest IP address. The DR originates PIM register messages for the directly connected multicast sources and sends PIM join messages toward the rendezvous point (RP) for directly connected receivers. Values range from 1 to 4294967295. The default is 1.
Hello authentication mode	Enables an MD5 hash authentication key, or password, in PIM hello messages on the interface so that directly connected neighbors can authenticate each other. The PIM hello messages are IPsec encoded using the Authentication Header (AH) option. You can enter an unencrypted (cleartext) key or one of these values followed by a space and the MD5 authentication key: • 0—Specifies an unencrypted (cleartext) key • 3—Specifies a 3-DES encrypted key • 7—Specifies a Cisco Type 7 encrypted key The authentication key can be up to 16 characters. The default is disabled. **Note** PIM6 does not support hello authentication.
Hello interval	Configures the interval at which hello messages are sent in milliseconds. The range is from 1000 to 18724286. The default is 30000. **Note** See the *Cisco Nexus 7000 Series NX-OS Verified Scalability Guide* for the verified range of this parameter and associated PIM neighbor scale.
Domain border	Enables the interface to be on the border of a PIM domain so that no bootstrap, candidate-RP, or Auto-RP messages are sent or received on the interface. The default is disabled. **Note** PIM6 does not support the Auto-RP method.

Parameter	Description
Neighbor policy	Configures which PIM neighbors to become adjacent to based on a route-map policy[3] where you can specify IP addresses to become adjacent to with the **match ip[v6] address** command. If the policy name does not exist, or no IP addresses are configured in a policy, adjacency is established with all neighbors. The default is to become adjacent with all PIM neighbors. **Note** We recommend that you should configure this feature only if you are an experienced network administrator.

[3] To configure route-map policies, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

Configuring PIM Sparse Mode Parameters

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip pim auto-rp** {**listen** [**forward**] **forward** [**listen**]}
	Example:
	`switch(config)# ip pim auto-rp listen`
		(Optional) Enables listening or forwarding of Auto-RP messages. The default is disabled, which means that the software does not listen to or forward Auto-RP messages.

Step 3	**ip pim bsr** {**listen** [**forward**] **forward** [**listen**]}
	Example:
	`switch(config)# ip pim bsr forward`
		(Optional) Enables listening or forwarding of Auto-RP messages. The default is disabled, which means that the software does not listen to or forward Auto-RP messages.

Step 4	**show ip pim rp** [*ip-prefix*] [**forward** | *vrf* [**vrf-name** | **all**] **Example:** `switch(config)# show ip pim rp`	(Optional) Enables listening or forwarding of Auto-RP messages. The default is disabled, which means that the software does not listen to or forward Auto-RP messages.
Step 5	**ip pim bidir-rp-limit** *limit* **Example:** `switch(config)# ip pim bidir-rp-limit 4`	(Optional) Specifies the number of Bidir RPs that you can configure for IPv4. The maximum number of Bidir RPs supported per VRF for PIM and PIM6 combined cannot exceed 8. Values range from 0 to 8. The default is 6.

	Command or Action	Purpose
Step 6	**ip pim register-rate-limit** *rate* **Example:** `switch(config)# ip pim register-rate-limit 1000`	(Optional) Configures the rate limit in packets per second. The range is from 1 to 65,535. The default is no limit.
Step 7	[**ip** | **ipv4**] **routing multicast holddown** *holddown-period* **Example:** `switch(config)# ip routing multicast holddown 100`	(Optional) Configures the initial holddown period in seconds. The range is from 90 to 210. Specify 0 to disable the holddown period. The default is 210.
Step 8	**show running-configuration pim** **Example:** `switch(config)# show running-configuration pim`	(Optional) Displays PIM running-configuration information, including the Bidir RP limit and register rate limit.

Step 9	**interface** *interface*
	Example:
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enters interface mode on the interface type and number, such as ethernet slot/port.

Step 10	**ip pim sparse-mode** **Example:** `switch(config-if)# ip pim sparse-mode`	Enables PIM sparse mode on this interface. The default is disabled.
Step 11	**ip pim dr-priority** *priority* **Example:** `switch(config-if)# ip pim dr-priority 192`	(Optional) Sets the designated router (DR) priority that is advertised in PIM hello messages. Values range from 1 to 4294967295. The default is 1.
Step 12	**ip pim hello-authentication ah-md5** *auth-key* **Example:** `switch(config-if)# ip pim hello-authentication ah-md5 my_key`	(Optional) Enables an MD5 hash authentication key in PIM hello messages. You can enter an unencrypted (cleartext) key or one of these values followed by a space and the MD5 authentication key: • 0—Specifies an unencrypted (cleartext) key • 3—Specifies a 3-DES encrypted key • 7—Specifies a Cisco Type 7 encrypted key The key can be up to 16 characters. The default is disabled.
Step 13	**ip pim hello-interval** *interval* **Example:**	(Optional) Configures the interval at which hello messages are sent in milliseconds. The

	Command or Action	Purpose
	`switch(config-if)# ip pim hello-interval 25000`	range is from 1000 to 18724286. The default is 30000. **Note** Before Cisco NX-OS Release 5.2(1), the minimum value was 1 millisecond.
Step 14	**ip pim border** **Example:** `switch(config-if)# ip pim border`	(Optional) Enables the interface to be on the border of a PIM domain so that no bootstrap, candidate-RP, or Auto-RP messages are sent or received on the interface. The default is disabled. **Note** When you use **ip pim border** command, the PIM border starts to work as a first-hop router under certain conditions. For information about PIM Multicast Border Router, see RFC 4601.
Step 15	**ip pim neighbor-policy** *policy-name* **Example:** `switch(config-if)# ip pim neighbor-policy my_neighbor_policy`	(Optional) Enables the interface to be on the border of a PIM domain so that no bootstrap, candidate-RP, or Auto-RP messages are sent or received on the interface. The default is disabled. (Optional) Configures which PIM neighbors to become adjacent to based on a route-map policy with the **match ip address** command. The policy name can be up to 63 characters. The default is to become adjacent with all PIM neighbors. **Note** We recommend that you should configure this feature only if you are an experienced network administrator.
Step 16	**show ip pim interface** [*interface* | **brief**] [**vrf** [*vrf-name* | **all**] **Example:** `switch(config-if)# show ip pim interface`	(Optional) Displays PIM interface information.
Step 17	**copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	(Optional) Saves configuration changes. (Optional) Configures which PIM neighbors to become adjacent to based on a route-map policy with the **match ip address** command. The policy name can be up to 63 characters. The default is to become adjacent with all PIM neighbors.

Configuring PIM6 Sparse Mode Parameters

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ipv6 pim bsr**{*listen*[*forward*] | *forward*[*listen*]}
	Example:
	`switch(config)# ip pim auto-rp listen`
		(Optional) Enables listening or forwarding of BSR messages. The default is disabled, which means that the software does not listen or forward BSR messages.

Step 3	**show ipv6 pim rp** [*ipv6-prefix*][**vrf***vrf-name*|**all**] **Example:** `switch(config)# show ipv6 pim rp`	(Optional) Displays PIM6 RP information, including BSR listen and forward states.
Step 4	**ipv6 pim bidir-rp-limit** *limit* **Example:** `switch(config)# ipv6 pim bidir-rp-limit 4`	(Optional) Specifies the number of Bidir RPs that you can configure for IPv6. The maximum number of Bidir RPs supported per VRF for PIM and PIM6 combined cannot exceed 8. Values range from 0 to 8. The default is 2.
Step 5	**ipv6 pim register-rate-limit** *rate* **Example:** `switch(config)# ipv6 pim register-rate-limit 1000`	(Optional) Configures the rate limit in packets per second. The range is from 1 to 65,535. The default is no limit.
Step 6	**ipv6 routing multicast holddown** *holddown-period* **Example:** `switch(config)# ipv6 routing multicast holddown 100`	(Optional) Configures the initial holddown period in seconds. The range is from 90 to 210. Specify 0 to disable the holddown period. The default is 210.
Step 7	**show running-configuration pim6** **Example:** `switch(config)# show running-configuration pim6`	(Optional) Displays PIM6 running-configuration information, including the Bidir RP limit and register rate limit.

Step 8	**interface** *interface*
	Example:
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enters interface mode on the specified interface.

	Command or Action	Purpose
Step 9	**ipv6 pim sparse-mode** **Example:** `switch(config-if)# ipv6 pim sparse-mode`	Enables PIM sparse mode on this interface. The default is disabled.
Step 10	**ipv6 pim dr-priority** *priority* **Example:** `switch(config-if)# ipv6 pim dr-priority 192`	(Optional) Sets the designated router (DR) priority that is advertised in PIM6 hello messages. Values range from 1 to 4294967295. The default is 1.
Step 11	**ipv6 pim hello-interval** *interval* **Example:** `switch(config-if)# ipv6 pim hello-interval 25000`	(Optional) Configures the interval at which hello messages are sent in milliseconds. The range is from 1000 to 18724286. The default is 30000. **Note** Before Cisco NX-OS Release 5.2(1), the minimum value was 1 millisecond.
Step 12	**ipv6 pim border** **Example:** `switch(config-if)# ipv6 pim border`	(Optional) Enables the interface to be on the border of a PIM6 domain so that no bootstrap, candidate-RP, or Auto-RP messages are sent or received on the interface. The default is disabled. **Note** Before Cisco NX-OS Release 5.2(1), the minimum value was 1 millisecond.
Step 13	**ipv6 pim neighbor-policy** *policy-name* **Example:** `switch(config-if)# ip pim border`	(Optional) Configures which PIM6 neighbors to become adjacent to based on a route-map policy with the **match ipv6 address** command. The policy name can be up to 63 characters. The default is to become adjacent with all PIM6 neighbors. **Note** We recommend that you should configure this feature only if you are an experienced network administrator.
Step 14	**show ipv6 pim interface** [*interface* | brief] [**vrf**vrf-name |**all**] **Example:** `switch(config-if)# show ipv6 pim interface`	(Optional) Displays PIM6 interface information.
Step 15	**copy running-config startup-config** **Example:**	(Optional) Saves configuration changes.

	Command or Action	Purpose
	`switch(config-if)# copy running-config startup-config`	

IGMP Querier

IGMP Querier Overview

The IGMP Querier feature supports the sending of Internet Group Management Protocol (IGMP) queries from a router only if the router is a multicast (PIM)-enabled) router. IGMP router functionality will be enabled only when PIM is enabled on the interface. IGMP router functionality will be disabled when PIM is disabled on the interface. If IGMP router functionality is enabled and PIM is disabled subsequently, then the router functionality will be disabled.

Enabling IGMP Querier

Perform this task to enable the sending of IGMP queries from a router only if the router is a multicast (PIM-enabled) router.

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enters global configuration mode.
Step 2	**interface** *type number* **Example:** `switch(config)# interface Ethernet 0/0`	Specifies an interface type and number, and places the device in interface configuration mode.
Step 3	**ip pim sparse-mode**] **Example:** `switch(config-if)# ip pim sparse-mode`	Enables PIM sparse-mode on an interface.

Step 4	**end**
	Example:
	`switch(config-if)# exit`
		Enter this command to go to privileged EXEC mode.

Step 5	**show ip igmp interface** **Example:** `switch# show ip igmp interface`	(Optional) Displays multicast-related information (including information on the IGMP querier) for an interface.

Example: Enabling IGMP Querier

The following example shows how to enable IGMP Querier:

```
switch# configure terminal
switch(config)# interface ethernet 2/1
switch(config-if)# ip pim sparse-mode
switch(config-if)#end
switch# show ip igmp interface

IGMP Interfaces for VRF "default", count: 2 Ethernet2/1, Interface status:
protocol-up/link-up/admin-up
  IP address: 10.11.11.1, IP subnet: 10.11.11.0/24
  Active querier: 10.11.11.1, version: 2, next query sent in: 00:01:57
  Membership count: 1
.
.
```

Configuring ASM and Bidir

Any Source Multicast (ASM) and bidirectional shared trees (Bidir) are multicast distribution modes that require the use of RPs to act as a shared root between sources and receivers of multicast data.

To configure ASM or Bidir mode, you configure sparse mode and the RP selection method, where you indicate the distribution mode and assign the range of multicast groups.

Note Bidir mode is not supported for vPCs. For more information about vPCs, see the *Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide*.

Configuring Static RPs

You can configure an RP statically by configuring the RP address on every router that participates in the PIM domain.

Note We recommend that the RP address uses the loopback interface and also the interface with the RP address must have **ip pim sparse-mode** enabled.

You can specify a route-map policy name that lists the group prefixes to use with the **match ip multicast** command.

Beginning with Cisco NX-OS Release 5.1(3), the **ip pim rp-address** command has been enhanced with the following functionalities:

- Added the prefix-list method of configuration in addition to existing route-map method.
- Added support for policy actions (route-map or prefix-list).

Note Cisco NX-OS always uses the longest-match prefix to find the RP. So, the behavior is the same irrespective of the position of the group prefix in the route map or in the prefix list.

The following example configuration produces the same output using Cisco NX-OS (231.1.1.0/24 is always denied irrespective of the sequence number):

```
ip prefix-list plist seq 10 deny 231.1.1.0/24
ip prefix-list plist seq 20 permit 231.1.0.0/16
ip prefix-list plist seq 10 permit 231.1.0.0/16
ip prefix-list plist seq 20 deny 231.1.1.0/24
```

This behavior differs from Cisco IOS. See the *Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*, behavior for more samples for the **ip pim rp-address** command.

Configuring Static RPs (PIM)

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ipv6 pim rp-address** *rp-address* [**group-list** *ipv6-prefix* | **route-map** *policy-name*] [**bidir**]
	Example:
	`switch(config)# ip pim rp-address 192.0.2.33 group-list 224.0.0.0/9`
	Example:
	`switch(config)# ip pim rp-address 192.0.2.34 group-list 224.128.0.0/9 bidir`
		Configures a PIM6 static RP address for a multicast group range. You can specify a route-map policy name that lists the group prefixes to use with the **match ip multicast** command. The mode is ASM unless you specify the **bidir** keyword. The default group range is ff00::0/8.
		Example 1 configures PIM6 ASM mode for the specified group range.
		Example 2 configures PIM6 Bidir mode for the specified group range.

Step 3	**show ip pim group-range** *ipv6-prefix*| **vrf** *vrf-name* **all** **Example:** `switch(config)# show ip pim group-range`	(Optional) Displays PIM6 RP information, including BSR listen and forward states.
Step 4	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring Static RPs (PIM6)

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ipv6 pim rp-address** *rp-address* [**group-list** *ipv6-prefix* | **route-map** *policy-nsmr*] [**bidir**]
	Example:
	`switch(config)# ipv6 pim rp-address 2001:0db8:0:abcd::1 group-list ff1e:abcd:def1::0/24`
	Example:
	`switch(config)# ipv6 pim rp-address 2001:0db8:0:abcd::2 group-list ff1e:abcd:def2::0/96 bidir`
		Configures a PIM6 static RP address for a multicast group range. You can specify a route-map policy name that lists the group prefixes to use with the **match ip multicast** command. The mode is ASM unless you specify the **bidir** keyword. The default group range is ff00::0/8.
		Example 1 configures PIM6 ASM mode for the specified group range.
		Example 2 configures PIM6 Bidir mode for the specified group range.

Step 3	**show ipv6 pim rp** *ipv6-prefix*|**vrf***vrf-name***all** **Example:** `switch(config)# show ipv6 pim group-range`	(Optional) Displays PIM6 modes and group ranges.
Step 4	**copy running-config startup-config** **Example:** `switch(config)# show ipv6 pim group-range`	(Optional) Displays PIM6 modes and group ranges.

Configuring BSRs

You configure BSRs by selecting candidate BSRs and RPs.

Caution Do not configure both Auto-RP and BSR protocols in the same network.

You can configure a candidate BSR with the arguments described on the Table below.

Table 14: Candidate BSR Arguments

Argument	Description
interface	Interface type and number used to derive the BSR source IP address used in bootstrap messages.

Argument	Description
hash-length	Number of high order 1s used to form a mask that is ANDed with group address ranges of candidate RPs to form a hash value. The mask determines the number of consecutive addresses to assign across RPs with the same group range. For PIM, this value ranges from 0 to 32 and has a default of 30. For PIM6, this value ranges from 0 to 128 and has a default of 126.
priority	Priority assigned to this BSR. The software elects the BSR with the highest priority, or if the BSR priorities match, the software elects the BSR with the highest IP address. This value ranges from 0, the lowest priority, to 255 and has a default of 64.

You can configure a candidate RP with the arguments and keywords described on this Table.

Table 15: BSR Candidate RP Arguments and Keywords

Argument or Keyword	Description
interface	Interface type and number used to derive the BSR source IP address used in Bootstrap messages.
group-list *ip-prefix*	Multicast groups handled by this RP specified in a prefix format.
interval	Number of seconds between sending candidate-RP messages. This value ranges from 1 to 65,535 and has a default of 60 seconds. **Note** We recommend that you configure the candidate RP interval to a minimum of 15 seconds.
priority	Priority assigned to this RP. The software elects the RP with the highest priority for a range of groups, or if the priorities match, the highest IP address. (The highest priority is the lowest numerical value.) This value ranges from 0, the highest priority, to 255 and has a default of 192. **Note** This priority differs from the BSR BSR-candidate priority, which prefers the highest value between 0 and 255.
bidir	Unless you specify bidir, this RP will be in ASM mode. If you specify bidir, the RP will be in Bidir mode.
route-map *policy-name*	Route-map policy name that defines the group prefixes where this feature is applied.

Tip You should choose the candidate BSRs and candidate RPs that have good connectivity to all parts of the PIM domain.

You can configure the same router to be both a BSR and a candidate RP. In a domain with many routers, you can select multiple candidate BSRs and RPs to automatically fail over to alternates if a BSR or an RP fails.

To configure candidate BSRs and RPs, follow these steps:

1. Configure whether each router in the PIM domain should listen to and forward BSR messages. A router configured as either a candidate RP or a candidate BSR automatically listens to and forwards all bootstrap router protocol messages, unless an interface is configured with the domain border feature. For more information, see the *Configuring PIM or PIM6 Sparse Mode*.
2. Select the routers to act as candidate BSRs and RPs.
3. Configure each candidate BSR and candidate RP as described in this section.
4. Configure BSR message filtering. See *Configuring Message Filtering*.

Configuring BSRs (PIM)

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip pim bsr listen forward** *listen|forward* | *forward|listen* **Example:** `switch(config)# ip pim bsr listen forward`	Configures listen and forward. Ensure that you have entered this command in each VRF on the remote PE.
Step 3	**ip pim bsr**[**bsr-candidate**] *interface* [**hash-len** *hash-length*] [**priorty** *priority]* **Example:** `switch(config)# ip pim bsr-candidate ethernet 2/1 hash-len 24`	Configures a candidate bootstrap router (BSR). The source IP address used in a bootstrap message is the IP address of the interface. The hash length ranges from 0 to 32 and has a default of 30. The priority ranges from 0 to 255 and has a default of 64.
Step 4	**ip pim sparse-mode** **Example:** `switch(config)# ip pim sparse-mode`	Enables PIM sparse mode on this interface. The default is disabled.
Step 5	**ip** [**bsr**] **rp-candidate** *interface* **group-list** *ip-prefix* **route-map** *policy-name* **priority** *priority* **interval** *interval* **bidir**	(Optional) Specifies the number of Bidir RPs that you can configure for IPv6. The maximum number of Bidir RPs supported per VRF for

	Command or Action	Purpose

	Example:
	`switch(config)# ip pim rp-candidate ethernet 2/1 group-list 239.0.0.0/24`
	Example:
	`switch(config)# ip pim rp-candidate ethernet 2/1 group-list 239.0.0.0/24 bidir`
		PIM and PIM6 combined cannot exceed 8. Values range from 0 to 8. The default is 2.
		Configures a candidate RP for BSR. The priority ranges from 0, the highest priority, to 65,535 and has a default of 192. The interval ranges from 1 to 65,535 seconds and has a default of 60.
		Note We recommend that you configure the candidate RP interval to a minimum of 15 seconds.
		Example 1 configures an ASM candidate RP.
		Example 2 configures a Bidir candidate RP.

Step 6	**show ip pim group-range** *ip-prefix* **vrf** *vrf-name* **all** **Example:** `switch(config)# show ip pim group-range`	(Optional) Displays PIM modes and group ranges.
Step 7	**ipv6 routing multicast holddown** *holddown-period* **Example:** `switch(config)# ipv6 routing multicast holddown 100`	(Optional) Saves configuration changes.

Configuring BSRs (PIM6)

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip pim [bsr] bsr-candidate** | *interface* [**hash-len** *hash-length* **priority** *priority]* **Example:** `switch(config)# ipv6 pim bsr-candidate ethernet 2/1 hash-len 24 priority 192`	Configures a candidate bootstrap router (BSR). The source IP address used in a bootstrap message is the IP address of the interface. The hash length ranges from 0 to 128 and has a default of 126. The priority ranges from 0, the lowest priority, to 255 and has a default of 64.
Step 3	**ipv6** [**bsr**] **rp-candidate** *interface* **group-list** *ipv6-prefix* [**route-map** *policy-name]* **priority** *priority* **interval** *interval* **bidir**]	Configures a candidate RP for BSR. The priority ranges from 0, the highest priority, to 65,535 and has a default of 192. The interval

	Command or Action	Purpose

	Example:
	`switch(config)# ipv6 pim rp-candidate ethernet 2/1 group-list ff1e:abcd:def1::0/24`
	Example:
	`switch(config)# ipv6 pim rp-candidate ethernet 2/1 group-list ff1e:abcd:def2::0/24 bidir`
		ranges from 1 to 65,535 seconds and has a default of 60.
		Example 1 configures an ASM candidate RP.
		Example 2 configures a Bidir candidate RP.

Step 4	**show ipv6 pim group-range** *ipv6-prefix* **vrf** *vrf-name* **all** **Example:** `switch(config)# show ipv6 pim group-range`	(Optional) Displays PIM6 modes and group ranges.
Step 5	**copy running-config startup-config** *holddown-period* **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring Auto-RP

You can configure Auto-RP by selecting candidate mapping agents and RPs. You can configure the same router to be both a mapping agent and a candidate RP.

Note Auto-RP is not supported by PIM6.

Caution Do not configure both Auto-RP and BSR protocols in the same network.

You can configure an Auto-RP mapping agent with the arguments described on this Table.

Table 16: Auto-RP Mapping Agent Arguments

Argument	Description
interface	Interface type and number used to derive the IP address of the Auto-RP mapping agent used in bootstrap messages.
scope *ttl*	Time-To-Live (TTL) value that represents the maximum number of hops that RP-Discovery messages are forwarded. This value can range from 1 to 255 and has a default of 32. **Note** See the border domain feature in the *Configuring PIM or PIM6 Sparse Mode*.

If you configure multiple Auto-RP mapping agents, only one is elected as the mapping agent for the domain. The elected mapping agent ensures that all candidate RP messages are sent out. All mapping agents receive the candidate RP messages and advertise the same RP cache in their RP-discovery messages.

You can configure a candidate RP with the arguments and keywords described on this Table.

Table 17: Auto-RP Candidate RP Arguments and Keywords

Argument or Keyword	Description
interface	Interface type and number used to derive the IP address of the candidate RP used in Bootstrap messages.
group-list *ip-prefix*	Multicast groups handled by this RP. It is specified in a prefix format.
scope *ttl*	Time-To-Live (TTL) value that represents the maximum number of hops that RP-Discovery messages are forwarded. This value can range from 1 to 255 and has a default of 32. **Note** See the border domain feature in the *Configuring PIM or PIM6 Sparse Mode*.
interval	Number of seconds between sending RP-Announce messages. This value can range from 1 to 65,535 and has a default of 60. **Note** We recommend that you configure the candidate RP interval to a minimum of 15 seconds.
bidir	If not specified, this RP will be in ASM mode. If specified, this RP will be in Bidir mode.
route-map *policy-name*	Route-map policy name that defines the group prefixes where this feature is applied.

Tip You should choose mapping agents and candidate RPs that have good connectivity to all parts of the PIM domain.

To configure Auto-RP mapping agents and candidate RPs, follow these steps:

1. For each router in the PIM domain, configure whether that router should listen and forward Auto-RP messages. A router configured as either a candidate RP or an Auto-RP mapping agent will automatically listen to and forward all Auto-RP protocol messages, unless an interface is configured with the domain border feature. For more information, see the *Configuring PIM or PIM6 Sparse Mode*.
2. Select the routers to act as mapping agents and candidate RPs.
3. Configure each mapping agent and candidate RP as described in this section.

4. Configure Auto-RP message filtering. See *Configuring Message Filtering*.

Ensure that you have installed the Enterprise Services license and enabled PIM or PIM6.

Configuring Auto RP (PIM)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM or PIM6.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip pim** {**send-rp-discovery** | { **auto-rp mapping-agent** }} *interface* [**scope** *ttl*] **Example:** `sswitch(config)# ip pim auto-rp mapping-agent ethernet 2/1`	Configures an Auto-RP mapping agent. The source IP address used in Auto-RP Discovery messages is the IP address of the interface. The default scope is 32.

Step 3	**ip pim** { |**send-rp-announce** | {**auto-rp rp-candidate**]}**auto***interface* {**group-list** *ip-prefix* | **route_map** *policy-name*} [**scope** *ttl*] **interval** *interval*] [**bidir**
	Example:
	`switch(config)# ip pim auto-rp rp-candidate ethernet 2/1 group-list 239.0.0.0/24`
	Example:
	`switch(config)# ip pim auto-rp rp-candidate ethernet 2/1 group-list 239.0.0.0/24 bidir`
		Configures an Auto-RP candidate RP. The default scope is 32. The default interval is 60 seconds. By default, the command creates an ASM candidate RP. For parameter details, see Table 4-8.
		Note We recommend that you configure the candidate RP interval to a minimum of 15 seconds.
		Example1 configures an ASM candidate RP.
		Example 2 configures a Bidir candidate RP.

Step 4	**ip pim sparse-mode** **Example:** `switch(config)# ip pim sparse-mode`	Enables PIM sparse mode on the interface. The default is disabled.
Step 5	**show ip pim group-range** *lip-prefix*] **vrf** *vrf-name* | **all**] **Example:** `switch(config)# show ip pim group-range`	(Optional) Displays PIM modes and group ranges.
Step 6	**copy running-config startup-config** *rate* **Example:**	(Optional) Saves configuration changes.

	Command or Action	Purpose
	`switch(config)# copy running-config startup-config`	

Configuring a PIM Anycast-RP Set

To configure a PIM Anycast-RP set, follow these steps:

1. Select the routers in the PIM Anycast-RP set.
2. Select an IP address for the PIM Anycast-RP set.
3. Configure each peer RP in the PIM Anycast-RP set as described in this section.

Configuring a PIM Anycast RP Set (PIM)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM or PIM6.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**interface loopback** *number* **Example:** `switch(config)# interface loopback 0`	Configures an interface loopback. This example configures interface loopback 0.
Step 3	**ip address** *ip-prefix* **Example:** `switch(config-if)# ip address 192.0.2.3/32`	Configures an IP address for this interface. This example configures an IP address for the Anycast-RP.
Step 4	**ip pim sparse-mode**	Enables PIM.
Step 5	**exit** **Example:** `switch(config)# exit`	Returns to configuration mode.
Step 6	**ip pim anycast-rp** *anycast-rp-address anycast-rp-peer-address* **Example:** `switch(config)# ip pim anycast-rp 192.0.2.3 192.0.2.31`	Configures a PIM Anycast-RP peer address for the specified Anycast-RP address. Each command with the same Anycast-RP address forms an Anycast-RP set. The IP addresses of RPs are used for communication with RPs in the set.

	Command or Action	Purpose
Step 7	Repeat Step 5 using the same Anycast-RP-address for each RP in the RP set (including the local router).	--
Step 8	**show ip pim group-range** [*ip-prefix*] [**vrf** *vrf-name* | **all**] **Example:** `switch(config)# show ip pim group-range`	Configures a PIM Anycast-RP peer address for the specified Anycast-RP address. Each command with the same Anycast-RP address forms an Anycast-RP set. The IP addresses of RPs are used for communication with RPs in the set.
Step 9	**copy running-config startup-config** [*ip-prefix*] [**vrf** *vrf-name* | **all**] **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring a PIM Anycast RP Set (PIM6)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM or PIM6.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**interface loopback** *number* **Example:** `switch(config)# interface loopback 0`	Configures an interface loopback. This example configures interface loopback 0.
Step 3	**ipv6 address** *ipv6-prefix* **Example:** `switch(config-if)# ipv6 address 2001:0db8:0:abcd::3/32`	Configures an IP address for this interface. This example configures an IP address for the Anycast-RP.
Step 4	**ip pim sparse-mode**	Enables PIM.
Step 5	**exit** **Example:** `switch(config)# exit`	Returns to configuration mode.
Step 6	**ipv6 pim anycast-rp** *anycast-rp-address anycast-rp-peer-address*	Configures a PIM6 Anycast-RP peer address for the specified Anycast-RP address. Each

	Command or Action	Purpose
	Example: `switch(config)# ipv6 pim anycast-rp 2001:0db8:0:abcd::3 2001:0db8:0:abcd::31`	command with the same Anycast-RP address forms an Anycast-RP set. The IP addresses of RPs are used for communication with RPs in the set.
Step 7	Repeat Step 5 using the same Anycast-RP-address for each RP in the RP set (including the local router).	--
Step 8	**show ipv6 pim group-range** [*ipv6-prefix*] [**vrf** *vrf-name* | **all**] **Example:** `switch(config)# show ipv6 pim group-range`	(Optional) Displays PIM6 modes and group ranges.
Step 9	**copy running-config startup-config** [*ip-prefix*] [**vrf** *vrf-name* | **all**] **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring Shared Trees Only for ASM

You can configure shared trees only on the last-hop router for Any Source Multicast (ASM) groups, which means that the router never switches over from the shared tree to the SPT when a receiver joins an active group. You can specify a group range where the use of shared trees is to be enforced with the **match ip**[**v6**] **multicast** command. This option does not affect the normal operation of the router when a source tree join-prune message is received.

Note The Cisco NX-OS software does not support the shared-tree feature on vPCs. For more information about vPCs, see the *Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide, Release 6.x*

The default is disabled, which means that the software can switch over to source trees.

Note In ASM mode, only the last-hop router switches from the shared tree to the SPT.

Configuring Shared Trees Only for ASM (PIM)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip pim use-shared-tree-only group-list** *policy-name* **Example:** `switch(config)# ip pim use-shared-tree-only group-list my_group_policy`	Builds only shared trees, which means that the software never switches over from the shared tree to the SPT. You specify a route-map policy name that lists the groups to use with the **match ip multicast** command. By default, the software triggers a PIM (S, G) join toward the source when it receives multicast packets for a source for which it has the (*, G) state.
Step 3	**show ip pim group-range** [*ip-prefix*] **vrf** *vrf-name* | **all** **Example:** `switch(config)# show ip pim group-range`	(Optional) Displays PIM modes and group ranges.
Step 4	**copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring Shared Trees Only for ASM (PIM6)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled for PIM6.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ipv6 pim use-shared-tree-only group-list** *policy-name* **Example:** `switch(config)# ipv6 pim use-shared-tree-only group-list my_group_policy`	Builds only shared trees, which means that the software never switches over from the shared tree to the SPT. You specify a route-map policy name that lists the groups to use with the **match ipv6 multicast command**. By default, the software triggers a PIM (S, G) join toward the source when it receives multicast packets for a source for which it has the (*, G) state.

	Command or Action	Purpose
Step 3	**show ipv6 pim group-range** [*ip-prefix*] **vrf** *vrf-name* | **all** **Example:** `switch(config)# show ipv6 pim group-range`	(Optional) Displays PIM6 modes and group ranges.
Step 4	**copy running-config startup-config** **Example:** `switch(config-if)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring SSM

Source-Specific Multicast (SSM) is a multicast distribution mode where the software on the DR connected to a receiver that is requesting data for a multicast source builds a shortest path tree (SPT) to that source.

On an IPv4 network, a host can request multicast data for a specific source only if it is running IGMPv3 and the DR for that host is running IGMPv3. You will usually enable IGMPv3 when you configure an interface for PIM in the SSM mode. For hosts running IGMPv1 or IGMPv2, you can configure a group to source mapping using SSM translation. For more information, see *Configuring IGMP* and *Configuring MLD*.

You can configure the group range that is used by SSM by specifying values on the command line. By default, the SSM group range for PIM is 232.0.0.0/8 and for PIM6 is FF3x/96.

You can specify a route-map policy name that lists the group prefixes to use with the **match ip multicast** command.

Note If you want to use the default SSM group range, you do not need to configure the SSM group range.

Configuring SSM (PIM)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**[no] ip pim ssm range** { *ip-prefix* | **none** | **route-map***policy-name* } **Example:**	Configures up to four group ranges to be treated in SSM mode. You can specify a route-map policy name that lists the group prefixes to use with the **match ip multicast** command. The

	Command or Action	Purpose

	`switch(config)# ip pim ssm range 239.128.1.0/24`
	Example:
	`switch(config)# no ip pim ssm range none`
		default range is 232.0.0.0/8. If the keyword **none** is specified, all group ranges are removed.
		The **no** option removes the specified prefix from the SSM range, or removes the route-map policy. If the keyword **none** is specified, resets the SSM range to the default of 232.0.0.0/8.

Step 3	**show ip pim group-range** [*ip-prefix*] **vrf** *vrf-name* | **all**] **Example:** `switch(config)# show ip pim group-range`	(Optional) Displays PIM mode and group ranges.
Step 4	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring SSM (PIM6)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM6.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**[no] ipv6 pim ssm** range { *ipv6-prefix* | **none** | **route-map** *policy-name* }
	Example:
	`switch(config)# ipv6 pim ssm range FF30::0/32`
		Configures up to four group ranges to be treated in SSM mode. You can specify a route-map policy name that lists the group prefixes to use with the match ip multicast command. If the keyword none is specified, all group ranges are removed. The default range is FF3x/96.

Step 3	**show ipv6 pim group-range** [*ipv6-prefix*] **vrf***vrf-name* | **all**]	(Optional) Displays PIM6 modes and group ranges.
Step 4	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring RPF Routes for Multicast

You can define RPF routes for multicast when you want multicast data to diverge from the unicast traffic path. You can define RPF routes for multicast on border routers to enable reverse path forwarding (RPF) to an external network.

Multicast routes are used not to directly forward traffic but to make RPF checks. RPF routes for multicast cannot be redistributed.

Note IPv6 static multicast routes are not supported.

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM or PIM6.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip mroute**{*ip-addr mask* | *ip-prefix*} {*next-hop* | *nh-prefix* | *interface*} [*route-preference*] [**vrf** *vrf-name*] **Example:** `switch(config)# ip mroute 192.0.2.33/1 224.0.0.0/1`	Configures an RPF route for multicast for use in RPF calculations. Route preference values range from 1 to 255. The default preference is 1.
Step 3	**show ip static-route** [**multicast**] [**vrf** *vrf-name*] **Example:** `switch(config)# show ip static-route multicast`	(Optional) Displays configured static routes.
Step 4	**copy running-config startup-config** [*ip-prefix*] **vrf***vrf-name* | **all**	(Optional) Saves configuration changes.

Disabling Multicast Multipath

By default, the RPF interface for multicast is chosen automatically when there are multiple ECMP paths available. Disabling the automatic selection allows you to specify a single RPF interface for multicast.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip multicast multipath none** **Example:** `switch(config)# ip multicast multipath none`	Disables multicast multipath.
Step 3	**clear ip mroute * vrf** *vrf-name* **Example:** `switch(config)# clear ip mroute *`	Clears multipath routes and activates multicast multipath suppression.

Enabling ECMP Multicast Load Splitting Based on Source Group and Next-Hop Address

Perform this task to enable ECMP multicast load splitting of multicast traffic based on source, group, and next-hop address (using the next-hop-based S-G-hash algorithm) to take advantage of multiple paths through the network. The next-hop-based S-G-hash algorithm is predictable because no randomization is used in calculating the hash value. Unlike the S-hash and basic S-G-hash algorithms, the hash mechanism used by the next-hop-based S-G-hash algorithm is not subject to polarization.

The next-hop-based S-G-hash algorithm provides more flexible support for ECMP multicast load splitting than S-hash algorithm and eliminates the polarization problem. Using the next-hop-based S-G-hash algorithm for ECMP multicast load splitting enables multicast traffic from devices that send many streams to groups or that broadcast many channels, such as IPTV servers or MPEG video servers, to be more effectively load split across equal-cost paths.

Procedure

	Command or Action	Purpose
Step 1	**configure terminal** **Example:** `switch# configure terminal`	Enters global configuration mode.
Step 2	**ip multicast multipath s-g-hash next-hop-based** **Example:** `switch(config)# ip multicast multipath s-g-hash next-hop-based`	Enables ECMP multicast load splitting based on source, group, and next-hop-address using the next-hop-based S-G-hash algorithm. • Because this command changes the way an RPF neighbor is selected, it must be configured consistently on all routers in a redundant topology to avoid looping.

	Command or Action	Purpose
		Note Be sure to enable the **ip multicast multipath** command on the router that is supposed to be the receiver for traffic from more than one incoming interfaces, which is opposite to unicast routing. From the perspective of unicast, multicast is active on the sending router connecting to more than one outgoing interfaces.
Step 3	Repeat Steps 1 through 3 on all the routers in a redundant topology.	--

Step 4	**end**
	Example:
	`switch(config)# end`
		Exits global configuration mode and returns to privileged EXEC mode.

Step 5	**show ip rpf** *source-address* [*group-address*] **Example:** `switch# show ip rpf 10.1.1.2`	(Optional) Displays the information that IP multicast routing uses to perform the RPF check. • Use this command to verify RPF selection so as to ensure that IP multicast traffic is being properly load split.
Step 6	**show ip route** *ip-address* **Example:** `switch# show ip route 10.1.1.2`	(Optional) Displays the current state of the IP routing table. • Use this command to verify that there multiple paths available to a source or RP for ECMP multicast load splitting. • For the *ip-address* argument, enter the IP address of a source to validate that there are multiple paths available to the source (for shortest path trees) or the IP address of an RP to validate that there are multiple paths available to the RP (for shared trees).

Example: Enabling ECMP Multicast Load Splitting Based on Source Group and Next-Hop Address

The following example shows how to enable ECMP multicast load splitting on a router based on source, group, and next-hop address using the next-hop-based S-G-hash algorithm:

```
switch(config)# ip multicast multipath s-g-hash next-hop-based
```

Configuring Route Maps to Control RP Information Distribution

You can configure route maps to help protect against some RP configuration errors and malicious attacks. You use route maps in commands that are described in the *Configuring Message Filtering*.

By configuring route maps, you can control distribution of RP information that is distributed throughout the network. You specify the BSRs or mapping agents to be listened to on each client router and the list of candidate RPs to be advertised (listened to) on each BSR and mapping agent to ensure that what is advertised is what you expect.

See the *Configuring BSRs* and *Configuring Auto-RP* for more information.

Note Only the **match ipv6 multicast** command has an effect in the route map.

Ensure that you have installed the Enterprise Services license and enabled PIM or PIM6.

Configuring Route Maps to Control RP Information Distribution (PIM)

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**route-map** *map-name* | **permit** | **deny**] [*sequence-number*]
	Example:
	`switch(config)# route-map ASM_only permit 10`
	`switch(config-route-map)#`
	Example:
	`switch(config)# route-map Bidir_only permit 10`
	`switch(config-route-map)#`
		Enters route-map configuration mode.
		Note This configuration method uses the **permit** keyword.

Step 3	**match ip multicast** {{**rp** *ip-address* [**rp-type** *rp-type*]} {{**group-range** {*gadrr_start* **to** *gadrr_end*} | {*group ip-prefix*}} {**source** *source-ip-address*}
	Example:
	`switch(config-route-map)# match ip multicast group 224.0.0.0/4 rp 0.0.0.0/0 rp-type ASM`
	Example:
		Matches the group, RP, and RP type specified. You can specify the RP type (ASM or Bidir). This configuration method requires the group and RP specified as shown in the examples.
		Note BSR RP, auto-RP, and static RP cannot use the **group-range** keyword. This command allows both permit or deny. Some match mask commands do not allow permit or deny.

	Command or Action	Purpose
	`switch(config-route-map)# match ip multicast group 224.0.0.0/4 rp 0.0.0.0/0 rp-type Bdir`	
Step 4	**show route-map** **Example:** `switch(config-route-map)# show route-map`	(Optional) Displays configured route maps.
Step 5	**copy running-config startup-config** **Example:** `switch(config-route-map)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring Route Maps to Control RP Information Distribution (PIM6)

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**route-map** *map-name* [**permit**| **deny**] [*sequence-number*]
	Example:
	`switch(config)# route-map ASM_only permit 10`
	`switch(config-route-map)#`
	Example:
	`switch(config)# route-map Bidir_only permit 10`
	`switch(config-route-map)#`
		Enters route-map configuration mode.
		Note This configuration method uses the **permit** keyword.

Step 3	**match ipv6 multicast** {{**rp** *ip-address* [**rp-type** *rp-type*]} {{**group-range** {*gadrr_start* **to** *gadrr_end*} | {*group ip-prefix*}} {**source** *source-ip-address*}
	Example:
	`switch(config-route-map)# match ip multicast group 224.0.0.0/4 rp 0.0.0.0/0 rp-type ASM`
	Example:
	`switch(config-route-map)# match ip multicast group 224.0.0.0/4 rp 0.0.0.0/0 rp-type Bdir`
		Matches the group, RP, and RP type specified. You can specify the RP type (ASM or Bidir). This configuration method requires the group and RP specified as shown in the examples.
		Note BSR RP, auto-RP, and static RP cannot use the **group-range** keyword. This command allows both permit or deny. Some match mask commands do not allow permit or deny.

	Command or Action	Purpose
Step 4	**show route-map** **Example:** `switch(config-route-map)# show route-map`	(Optional) Displays configured route maps.
Step 5	**copy running-config startup-config** **Example:** `switch(config-route-map)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring Message Filtering

Note Prefix matches in the rp-candidate-policy must be exact relative to what the c-rp is advertising. Subset matches are not possible.

You can configure filtering of the PIM and PIM6 messages described in the table below.

Table 18: PIM and PIM6 Message Filtering

Message Type	Description
Global to the Device	
Log Neighbor changes	Enables syslog messages that list the neighbor state changes to be generated. The default is disabled.
PIM register policy	Enables PIM register messages to be filtered based on a route-map policy[4] where you can specify group or group and source addresses with the **match ip[v6] multicast** command. This policy applies to routers that act as an RP. The default is disabled, which means that the software does not filter PIM register messages.
BSR candidate RP policy	Enables BSR candidate RP messages to be filtered by the router based on a route-map policy1 where you can specify the RP and group addresses and whether the type is Bidir or ASM with the **match ip[v6] multicast** command. This command can be used on routers that are eligible for BSR election. The default is no filtering of BSR messages.
BSR policy	Enables BSR messages to be filtered by the BSR client routers based on a route-map policy1 where you can specify BSR source addresses with the **match ip[v6] multicast** command. This command can be used on client routers that listen to BSR messages. The default is no filtering of BSR messages.

Message Type	Description
Auto-RP candidate RP policy	Enables Auto-RP announce messages to be filtered by the Auto-RP mapping agents based on a route-map policy1 where you can specify the RP and group addresses, and whether the type is Bidir or ASM with the **match ip multicast** command. This command can be used on a mapping agent. The default is no filtering of Auto-RP messages. **Note** PIM6 does not support the Auto-RP method.
Auto-RP mapping agent policy	Enables Auto-RP discover messages to be filtered by client routers based on a route-map policy1 where you can specify mapping agent source addresses with the **match ip multicast** command. This command can be used on client routers that listen to discover messages. The default is no filtering of Auto-RP messages. **Note** PIM6 does not support the Auto-RP method.
Per Device Interface	
Join-prune policy	Enables join-prune messages to be filtered based on a route-map policy1 where you can specify group, group and source, or group and RP addresses with the **match ip[v6] multicast** command. The default is no filtering of join-prune messages.

[4] For information about configuring route-map policies, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

Route maps as a filtering policy can be used (either **permit** or **deny** for each statement) for the following commands:

- **jp-policy** can use (S,G), (*,G), or (RP,G)
- **register-policy** can use (S,G) or (*,G)
- **igmp report-policy** can use (*,G) or (S,G)
- **state-limit reserver-policy** can use (*,G) or (S,G)
- **auto-rp rp-candidate-policy** can use (RP,G)
- **bsr rp-candidate-policy** can use (group-range/G, RP, RP-type)
- **autorp mapping-agent policy** can use (S)
- **bsr bsr-policy** can use (S)

Route maps as containers can be use for the following commands, where route-map action (**permit** or **deny**) is ignored:

- **ip pim rp-address route map** can use only G
- **ip pim ssm-range route map** can use only G
- **ip igmp static-oif route map** can use (S,G), (*,G), (S,G-range), (*,G-range)
- **ip igmp join-group route map** can use (S,G), (*,G), (S,G-range, (*, G-range)

Configuring Message Filtering (PIM)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled for PIM.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip pim log-neighbor-changes** **Example:** `switch(config)# ip pim log-neighbor-changes`	(Optional) Enables syslog messages that list the neighbor state changes to be generated. The default is disabled..
Step 3	**ip pim register-policy** *policy-name* **Example:** `switch(config)# ip pim register-policy my_register_policy`	(Optional) Enables PIM register messages to be filtered based on a route-map policy. You can specify group or group and source addresses with the **match ip multicast** command.
Step 4	**ip pim bsr rp-candidate-policy** *policy-name* **Example:** `switch(config)# ip pim bsr rp-candidate-policy my_bsr_rp_candidate_policy`	(Optional) Enables BSR candidate RP messages to be filtered by the router based on a route-map policy where you can specify the RP and group addresses and whether the type is Bidir or ASM with the **match ip multicast** command. This command can be used on routers that are eligible for BSR election. The default is no filtering of BSR messages.
Step 5	**ip pim bsr bsr-policy** *policy-name* **Example:** `switch(config)# ip pim bsr bsr-policy my_bsr_policy`	(Optional) Enables BSR messages to be filtered by the BSR client routers based on a route-map policy where you can specify BSR source addresses with the **match ip multicast** command. This command can be used on client routers that listen to BSR messages. The default is no filtering of BSR messages.

	Command or Action	Purpose
Step 6	**ip pim auto-rp rp-candidate-policy** *policy-name* **Example:** `switch(config)# ip pim auto-rp rp-candidate-policy my_auto_rp_candidate_policy`	(Optional) Enables Auto-RP announce messages to be filtered by the Auto-RP mapping agents based on a route-map policy where you can specify the RP and group addresses and whether the type is Bidir or ASM with the **match ip multicast** command. This command can be used on a mapping agent. The default is no filtering of Auto-RP messages.
Step 7	**ip pim auto-rp mapping-agent-policy** *policy-name* **Example:** `switch(config)# ip pim auto-rp mapping-agent-policy my_auto_rp_mapping_policy`	(Optional) Enables Auto-RP discover messages to be filtered by client routers based on a route-map policy where you can specify mapping agent source addresses with the **match ip multicast** command. This command can be used on client routers that listen to discover messages. The default is no filtering of Auto-RP messages.

Step 8	**interface** *interface*
	Example:
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enters interface mode on the specified interface.

Step 9	**ip pim jp-policy** *policy-name*[in | out] **Example:** `switch(config-if)# ip pim jp-policy my_jp_policy`	(Optional) Enables join-prune messages to be filtered based on a route-map policy where you can specify group, group and source, or group and RP addresses with the**match ip multicast** command. The default is no filtering of join-prune messages. Beginning with Cisco NX-OS Release 4.2(3), this command filters messages in both incoming and outgoing directions.
Step 10	**show run pim** **Example:** `switch(config-if)# show run pim`	(Optional) Displays PIM configuration commands.
Step 11	**copy running-config startup-config** *interval* **Example:** `switch(config-if)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring Message Filtering (PIM6)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled for PIM6.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ipv6 pim log-neighbor-changes** **Example:** `switch(config)# ipv6 pim log-neighbor-changes`	(Optional) Enables syslog messages that list the neighbor state changes to be generated. The default is disabled..
Step 3	**ipv6 pim register-policy** *policy-name* **Example:** `switch(config)# ipv6 pim register-policy my_register_policy`	(Optional) Enables PIM register messages to be filtered based on a route-map policy. You can specify group or group and source addresses with the **match ipv6 multicast** command. The default is disabled.
Step 4	**ipv6 pim bsr rp-candidate-policy** *policy-name* **Example:** `switch(config)# ipv6 pim bsr rp-candidate-policy my_bsr_rp_candidate_policy`	(Optional) Enables BSR candidate RP messages to be filtered by the router based on a route-map policy where you can specify the RP and group addresses and whether the type is Bidir or ASM with the **match ipv6 multicast** command. This command can be used on routers that are eligible for BSR election. The default is no filtering of BSR messages.
Step 5	**ipv6 pim bsr bsr-policy** *policy-name* **Example:** `switch(config)# ipv6 pim bsr bsr-policy my_bsr_policy`	(Optional) Enables BSR messages to be filtered by the BSR client routers based on a route-map policy where you can specify BSR source addresses with the **match ipv6 multicast** command. This command can be used on client routers that listen to BSR messages. The default is no filtering of BSR messages.

Step 6	**interface** *interface*
	Example:
	`switch(config)# interface ethernet 2/1`
	`switch(config-if)#`
		Enters interface mode on the specified interface.

Step 7	**ipv6 pim jp-policy** *policy-name*[**in** | **out**] **Example:** `switch(config-if)# ipv6 pim jp-policy my_jp_policy`	(Optional) Enables join-prune messages to be filtered based on a route-map policy where you can specify group, group and source, or group and RP addresses with the**match ipv6 multicast** command. The default is no filtering of join-prune messages. Beginning with Cisco NX-OS Release 4.2(3), this command filters messages in both incoming and outgoing directions.

	Command or Action	Purpose
Step 8	**show run pim6** **Example:** `switch(config-if)# show run pim6`	(Optional) Displays PIM6 configuration commands.
Step 9	**copy running-config startup-config** *interval* **Example:** `switch(config-if)# copy running-config startup-config`	(Optional) Saves configuration changes.

Restarting the PIM and PIM6 Processes

You can restart the PIM and PIM6 processes and optionally flush all routes. By default, routes are not flushed.

When routes are flushed, they are removed from the Multicast Routing Information Base (MRIB and M6RIB) and the Multicast Forwarding Information Base (MFIB and M6FIB).

When you restart PIM or PIM6, the following tasks are performed:

- The PIM database is deleted.
- The MRIB and MFIB are unaffected and forwarding of traffic continues.
- The multicast route ownership is verified through the MRIB.
- Periodic PIM join and prune messages from neighbors are used to repopulate the database.

Restarting the PIM Process (PIM)

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM.

Procedure

	Command or Action	Purpose
Step 1	**restart pim** **Example:** `switch# restart pim`	Restarts the PIM process.

Step 2	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 3	**ip pim flush-routes** **Example:** `switch(config)# ip pim flush-routes`	Removes routes when the PIM process is restarted. By default, routes are not flushed.

	Command or Action	Purpose
Step 4	**show running-configuration pim** **Example:** `switch(config)# show running-configuration pim`	(Optional) Displays the PIM running-configuration information, including the flush-routes command.
Step 5	**copy running-config startup-config** *policy-name* **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Restarting the PIM6 Process

Before you begin

Ensure that you have installed the Enterprise Services license and enabled PIM6.

Procedure

	Command or Action	Purpose
Step 1	**restart pim6** **Example:** `switch# restart pim`	Restarts the PIM process.

Step 2	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 3	**ipv6 pim flush-routes** **Example:** `switch(config)# ipv6 pim flush-routes`	Removes routes when the PIM6 process is restarted. By default, routes are not flushed.
Step 4	**show running-configuration pim6** **Example:** `switch(config)# show running-configuration pim6`	(Optional) Displays the PIM6 running-configuration information, including the **flush-routes** command.
Step 5	**copy running-config startup-config** *policy-name* **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring BFD for PIM in VRF Mode

Note You can configure BFD for PIM by either VRF or by interface.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**vrf context** *vrf-name*
	Example:
	`switch# vrf convrf-name`
	`text test`
	`switch(config-vrf)#`
		Enters VRF configuration mode.

Step 3	**ip pim bfd** **Example:** `switch(config-vrf)# ip pim bfd`	Enables BFD on the specified VRFs. **Note** You can also enter the **ip pim bfd** command in configuration mode, which enables BFD on VRF. Enters VRF configuration mode.

Configuring BFD for PIM in Interface Mode

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch(config)# interface ethernet 7/40`
	`switch(config-if)#`
		Enters global configuration mode.

Step 2	**interface** *interface-type*
	Example:
	`switch(config)# interface ethernet 7/40`
	`switch(config-if)#`
		Enters interface configuration mode.

Step 3	**config tip pim bfd instance** **Example:** `switch(config-if)# ip pim bfd instance`	Enables BFD on the specified interfaces. You can enable or disable BFD on RIM interfaces irrespective of whether BFD is enabled on the VRF.

	Command or Action	Purpose
Step 4	**exit** **Example:** `switch(config)# exit`	Exits out of VRF or interface configuration mode.
Step 5	**show running-configuration pim** **Example:** `switch(config)# show running-configuration pim`	(Optional) Displays the PIM running-configuration information.
Step 6	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Verifying the PIM and PIM6 Configuration

To display the PIM and PIM6 configurations information, perform one of the following tasks. Use the **show ip** form of the command for PIM and the **show ipv6** form of the command for PIM6.

Command	Description
show ip [**v6**] **mroute** {*source group* | *group* [*source*]} [**vrf** *vrf-name* | **all**]	Displays the IP or IPv6 multicast routing table.
show ip [**v6**] **pim df** [**vrf** *vrf-name* | **all**]	Displays the designated forwarder (DF) information for each RP by interface.
show ip [**v6**] **pim group-range** [**vrf** *vrf-name* | **all**]	Displays the learned or configured group ranges and modes. For similar information, see also the **show ip pim rp** command.
show ip [**v6**] **pim interface** [*interface* | **brief**] [**vrf** *vrf-name* | **all**]	Displays information by the interface.
show ip [**v6**] **pim neighbor** [**vrf** *vrf-name* | **all**]	Displays neighbors by the interface.
show ip [**v6**] **pim oif-list** *group* [*source*] [**vrf** *vrf-name* | **all**]	Displays all the interfaces in the OIF-list.
show ip [**v6**] **pim route** {source group | group [source]} [**vrf** *vrf-name* | **all**]	Displays information for each multicast route, including interfaces on which a PIM join for that (S, G) has been received.
show ip [**v6**] **pim rp** [**vrf** *vrf-name* | **all**]	Displays rendezvous points (RPs) known to the software, how they were learned, and their group ranges. For similar information, see also the **show ip pim group-range** command.

Command	Description
show ip [**v6**] **pim rp-hash** [**vrf** *vrf-name* | **all**]	Displays the bootstrap router (BSR) RP hash information. For information about the RP hash, see *RFC 5059*.
show running-configuration pim[**6**]	Displays the running-configuration information.
show startup-configuration pim[**6**]	Displays the startup-configuration information.
show ip [**v6**] **pim vrf** [*vrf-name* **| all**] [**detail**]	Displays per-VRF information.

For detailed information about the fields in the output from these commands, see the *Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*.

Displaying Statistics

You can display and clear PIM and PIM6 statistics by using the commands in this section.

Displaying PIM and PIM6 Statistics

You can display the PIM and PIM6 statistics and memory usage using the commands listed in the table below. Use the **show ip** form of the command for PIM and the **show ipv6** form of the command for PIM6.

Command	Description
show ip [v6] pim policy statistics	Displays policy statistics for Register, RP, and join-prune message policies.
show ip [v6] pim statistics [**vrf** *vrf-name* | **all**]	Displays global statistics. If PIM is in vPC mode, displays vPC statistics.

For detailed information about the fields in the output from these commands, see the *Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*.

Clearing PIM and PIM6 Statistics

You can clear the PIM and PIM6 statistics using the commands listed in the table below. Use the **show ip** form of the command for PIM and the **show ipv6** form of the command for PIM6.

Command	Description
ip [v6] pim interface statistics*interface*	Clears counters for the specified interface.
clear ip [v6] pim policy statistics	Clears policy counters for Register, RP, and join-prune message policies.
clear ip [v6] pim statistics [**vrf-name** | **all**]	Clears global counters handled by the PIM process.

Command	Description
clear ip mroute statistics {***** | *ipv4-grp-addr/prefix-length*} [**vrf** {*vrf-name* | **all** | **default** | **management**}]	Clears software and hardware statistics for all or specific multicast routes or multicast prefixes. **Note** Use the **show ip mroute** command to display the statistics for multicast route and prefixes.

Displaying Replication Engine Statistics

You can display replication engine statistics by using the **show hardware replication-engine statistics** [*module mod-no*] [*instance inst-no*] command.

Replication Engine Statistics Example

```
switch# show hard rep stat mod 10 inst 0
Replication Engine Statistics for Module 10 (N7K-M108X2-12L)

Instance 0 (ports 1-2):
Packet Counters:
Description                                  InPkts            OutPkts
-----------------------------------------------------------------------
Interface In Hi (port 1)                          0                  0
Interface In Lo (port 1)                          0                  0
Interface In Hi (port 2)                          0                  0
Interface In Lo (port 2)                          0                  0
Interface Out Hi (port 1)                         0                  0
Interface Out Lo (port 1)                         0                  0
Interface Out Hi (port 2)                         0                  0
Interface Out Lo (port 2)                         0                  0
Fabric In Hi                                      0                  0
Fabric In Lo                                      0                  0
Fabric Out Hi                                     0                  0
Fabric Out Lo                                     0                  0
Fwding Engine Tx                                  0                  0
Fwding Engine Rx                                  0                  0
Fwding Engine Ucast Rx                            0                  0
Fwding Engine Mcast Rx                            0                  0
Fwding Engine Rx                                  0                  0
Replication In Ucast                              0                  0
Replication Out Ucast                             0                  0
Replication In Mcast                              0                  0
Replication Out Mcast                             0                  0

Rates:
Description                         In PPS       In Bps       Out PPS      Out Bps
-----------------------------------------------------------------------------------
Interface In Hi (port 1)                 0            0             0            0
Interface In Lo (port 1)                 0            0             0            0
Interface In Hi (port 2)                 0            0             0            0
Interface In Lo (port 2)                 0            0             0            0
Interface Out Hi (port 1)                0            0             0            0
Interface Out Lo (port 1)                0            0             0            0
Interface Out Hi (port 2)                0            0             0            0
Interface Out Lo (port 2)                0            0             0            0
Fabric In Hi                             0            0             0            0
Fabric In Lo                             0            0             0            0
Fabric Out Hi                            0            0             0            0
Fabric Out Lo                            0            0             0            0
Fwding Engine Tx                         0            0             0            0
Fwding Engine Rx                         0            0             0            0
Fwding Engine Ucast Rx                   0            0             0            0
Fwding Engine Mcast Rx                   0            0             0            0
Fwding Engine Rx                         0            0             0            0
Replication In Ucast                     0            0             0            0
Replication Out Ucast                    0            0             0            0
Replication In Mcast                     0            0             0            0
Replication Out Mcast                    0            0             0            0


Drop Counters:
Description                                  Drops
---------------------------------------------------
Multicast/SPAN FIFO Drops                        0
SPAN Rate Limiter Drops                          0

SPAN Rate Limiter State: DISABLED

Peak Rates:
Packets per second:
Description                  Peak PPS             Date/Time
------------------------------------------------------------
Interface In (port 1)               0      yyyy/mm/dd hh:ss
Interface In (port 2)               0      yyyy/mm/dd hh:ss
Interface Out (port 1)              0      yyyy/mm/dd hh:ss
Interface Out (port 2)              0      yyyy/mm/dd hh:ss
Fabric In                           0      yyyy/mm/dd hh:ss
Fabric Out                          0      yyyy/mm/dd hh:ss
Replication In Ucast                0      yyyy/mm/dd hh:ss
Replication Out Ucast               0      yyyy/mm/dd hh:ss
Replication In Mcast                0      yyyy/mm/dd hh:ss
Replication Out Mcast               0      yyyy/mm/dd hh:ss

Bytes per second:
Description                  Peak Bps             Date/Time
------------------------------------------------------------
Interface In (port 1)               0      yyyy/mm/dd hh:ss
Interface In (port 2)               0      yyyy/mm/dd hh:ss
Interface Out (port 1)              0      yyyy/mm/dd hh:ss
Interface Out (port 2)              0      yyyy/mm/dd hh:ss
Fabric In                           0      yyyy/mm/dd hh:ss
Fabric Out                          0      yyyy/mm/dd hh:ss

switch#
```

Configuration Examples for PIM

Note See the *Multiple RPs Configured in a PIM Domain* for more configuration examples.

This section describes how to configure PIM using different data distribution modes and RP selection methods.

SSM Configuration Example

To configure PIM in SSM mode, follow these steps for each router in the PIM domain:

1. Configure PIM sparse mode parameters on the interfaces that you want to participate in the domain. We recommend that you enable PIM on all interfaces.

   ```
   switch# config t
   switch(config)# interface ethernet 2/1
   switch(config-if)# ip pim sparse-mode
   ```

2. Configure the parameters for IGMP that support SSM. See *Configuring IGMP* Usually, you configure IGMPv3 on PIM interfaces to support SSM.

   ```
   switch# config t
   switch(config)# interface ethernet 2/1
   switch(config-if)# ip igmp version 3
   ```

3. Configure the SSM range if you do not want to use the default range.

   ```
   switch# config t
   switch(config)# ip pim ssm range 239.128.1.0/24
   ```

4. Configure message filtering.

   ```
   switch# config t
   switch(config)# ip pim log-neighbor-changes
   ```

The following example shows how to configure PIM SSM mode:

```
config t
  interface ethernet 2/1
    ip pim sparse-mode
    ip igmp version 3
    exit
  ip pim ssm range 239.128.1.0/24
  ip pim log-neighbor-changes
```

BSR Configuration Example

To configure PIM in ASM mode using the BSR mechanism, follow these steps for each router in the PIM domain:

1. Configure PIM sparse mode parameters on the interfaces that you want to participate in the domain. We recommend that you enable PIM on all interfaces.

   ```
   switch# config t
   switch(config)# interface ethernet 2/1
   switch(config-if)# ip pim sparse-mode
   ```

2. Configure whether that router should listen and forward BSR messages

   ```
   switch# config t
   switch(config)# ip pim bsr forward listen
   ```

3. Configure the BSR parameters for each router that you want to act as a BSR.

```
switch# config t
switch(config)# ip pim bsr-candidate ethernet 2/1 hash-len 30
```

4. Configure the RP parameters for each router that you want to act as a candidate RP

```
switch# config t
switch(config)# ip pim rp-candidate ethernet 2/1 group-list 239.0.0.0/24
```

5. Configure message filtering.

```
switch# config t
switch(config)# ip pim log-neighbor-changes
```

This example shows how to configure PIM ASM mode using the BSR mechanism and how to configure the BSR and RP on the same router:

```
config t
  interface ethernet 2/1
    ip pim sparse-mode
    exit
  ip pim bsr forward listen
ip pim bsr-candidate ethernet 2/1 hash-len 30
  ip pim rp-candidate ethernet 2/1 group-list 239.0.0.0/24
  ip pim log-neighbor-changes
```

Auto-RP Configuration Example

To configure PIM in Bidir mode using the Auto-RP mechanism, follow these steps for each router in the PIM domain:

1. Configure PIM sparse mode parameters on the interfaces that you want to participate in the domain. We recommend that you enable PIM on all interfaces.

```
switch# config t
switch(config)# interface ethernet 2/1
switch(config-if)# ip pim sparse-mode
```

2. Configure whether that router should listen and forward Auto-RP messages.

```
switch# config t
switch(config)# ip pim auto-rp forward listen
```

3. Configure the mapping agent parameters for each router that you want to act as a mapping agent.

```
switch# config t
switch(config)# ip pim auto-rp mapping-agent ethernet 2/1
```

4. Configure the RP parameters for each router that you want to act as a candidate RP.

```
switch# config t
switch(config)# ip pim auto-rp rp-candidate ethernet 2/1 group-list 239.0.0.0/24 bidir
```

5. Configure message filtering.

```
switch# config t
switch(config)# ip pim log-neighbor-changes
```

This example shows how to configure PIM Bidir mode using the Auto-RP mechanism and how to configure the mapping agent and RP on the same router:

```
config t
  interface ethernet 2/1
    ip pim sparse-mode
    exit
  ip pim auto-rp listen
  ip pim auto-rp forward
  ip pim auto-rp mapping-agent ethernet 2/1
  ip pim auto-rp rp-candidate ethernet 2/1 group-list 239.0.0.0/24 bidir
  ip pim log-neighbor-changes
```

PIM Anycast RP Configuration Example

To configure ASM mode using the PIM Anycast-RP method, follow these steps for each router in the PIM domain:

1. Configure PIM sparse mode parameters on the interfaces that you want to participate in the domain. We recommend that you enable PIM on all interfaces.

```
switch# config t
switch(config)# interface ethernet 2/1
switch(config-if)# ip pim sparse-mode
```

2. Configure the RP address that you configure on all routers in the Anycast-RP set.

```
switch# config t
switch(config)# interface loopback 0
switch(config-if)# ip address 192.0.2.3/32
```

3. Configure a loopback with an address to use in communication between routers in the Anycast-RP set for each router that you want to be in the Anycast-RP set.

```
switch# config t
switch(config)# interface loopback 1
switch(config-if)# ip address 192.0.2.31/32
```

4. Configure the Anycast-RP parameters and repeat with the IP address of each Anycast-RP for each router that you want to be in the Anycast-RP set. This example shows two Anycast-RPs.

```
switch# config t
switch(config)# ip pim anycast-rp 192.0.2.3 193.0.2.31
switch(config)# ip pim anycast-rp 192.0.2.3 193.0.2.32
```

5. Configure message filtering.

```
switch# config t
switch(config)# ip pim log-neighbor-changes
```

This example shows how to configure PIM ASM mode using two Anycast-RPs:

```
config t
interface ethernet 2/1
ip pim sparse-mode
exit
interface loopback 0
```

```
ip address 192.0.2.3/32
exit
ip pim anycast-rp 192.0.2.3 192.0.2.31
ip pim anycast-rp 192.0.2.3 192.0.2.32
ip pim log-neighbor-changes
```

Prefix-Based and Route-Map-Based Configurations

```
ip prefix-list plist11 seq 10 deny 231.129.128.0/17
ip prefix-list plist11 seq 20 deny 231.129.0.0/16
ip prefix-list plist11 seq 30 deny 231.128.0.0/9
ip prefix-list plist11 seq 40 permit 231.0.0.0/8

ip prefix-list plist22 seq 10 deny 231.129.128.0/17
ip prefix-list plist22 seq 20 deny 231.129.0.0/16
ip prefix-list plist22 seq 30 permit 231.128.0.0/9
ip prefix-list plist22 seq 40 deny 231.0.0.0/8

ip prefix-list plist33 seq 10 deny 231.129.128.0/17
ip prefix-list plist33 seq 20 permit 231.129.0.0/16
ip prefix-list plist33 seq 30 deny 231.128.0.0/9
ip prefix-list plist33 seq 40 deny 231.0.0.0/8

ip pim rp-address 21.21.0.11 prefix-list plist11
ip pim rp-address 21.21.0.22 prefix-list plist22
ip pim rp-address 21.21.0.33 prefix-list plist33
route-map rmap11 deny 10
 match ip multicast group 231.129.128.0/17
route-map rmap11 deny 20
 match ip multicast group 231.129.0.0/16
route-map rmap11 deny 30
 match ip multicast group 231.128.0.0/9
route-map rmap11 permit 40
 match ip multicast group 231.0.0.0/8

route-map rmap22 deny 10
 match ip multicast group 231.129.128.0/17
route-map rmap22 deny 20
 match ip multicast group 231.129.0.0/16
route-map rmap22 permit 30
 match ip multicast group 231.128.0.0/9
route-map rmap22 deny 40
 match ip multicast group 231.0.0.0/8

route-map rmap33 deny 10
 match ip multicast group 231.129.128.0/17
route-map rmap33 permit 20
 match ip multicast group 231.129.0.0/16
route-map rmap33 deny 30
 match ip multicast group 231.128.0.0/9
route-map rmap33 deny 40
 match ip multicast group 231.0.0.0/8

ip pim rp-address 21.21.0.11 route-map rmap11
ip pim rp-address 21.21.0.22 route-map rmap22
ip pim rp-address 21.21.0.33 route-map rmap33
```

Output

```
dc3rtg-d2(config-if)# show ip pim rp
```

```
PIM RP Status Information for VRF "default"
BSR disabled
Auto-RP disabled
BSR RP Candidate policy: None
BSR RP policy: None
Auto-RP Announce policy: None
Auto-RP Discovery policy: None

RP: 21.21.0.11, (0), uptime: 00:12:36, expires: never,
  priority: 0, RP-source: (local), group-map: rmap11, group ranges:
      231.0.0.0/8  231.128.0.0/9 (deny)
      231.129.0.0/16 (deny) 231.129.128.0/17 (deny)
RP: 21.21.0.22, (0), uptime: 00:12:36, expires: never,
  priority: 0, RP-source: (local), group-map: rmap22, group ranges:
      231.0.0.0/8 (deny) 231.128.0.0/9
      231.129.0.0/16 (deny) 231.129.128.0/17 (deny)
RP: 21.21.0.33, (0), uptime: 00:12:36, expires: never,
  priority: 0, RP-source: (local), group-map: rmap33, group ranges:
      231.0.0.0/8 (deny) 231.128.0.0/9 (deny)
      231.129.0.0/16  231.129.128.0/17 (deny)

dc3rtg-d2(config-if)# show ip mroute
IP Multicast Routing Table for VRF "default"

(*, 231.1.1.1/32), uptime: 00:07:20, igmp pim ip
  Incoming interface: Ethernet2/1, RPF nbr: 1.1.0.1
  Outgoing interface list: (count: 1)
    loopback1, uptime: 00:07:20, igmp

(*, 231.128.1.1/32), uptime: 00:14:27, igmp pim ip
  Incoming interface: Ethernet2/1, RPF nbr: 1.1.0.1
  Outgoing interface list: (count: 1)
    loopback1, uptime: 00:14:27, igmp

(*, 231.129.1.1/32), uptime: 00:14:25, igmp pim ip
  Incoming interface: Ethernet2/1, RPF nbr: 1.1.0.1
  Outgoing interface list: (count: 1)
    loopback1, uptime: 00:14:25, igmp

(*, 231.129.128.1/32), uptime: 00:14:26, igmp pim ip
  Incoming interface: Null, RPF nbr: 0.0.0.0
  Outgoing interface list: (count: 1)
    loopback1, uptime: 00:14:26, igmp

(*, 232.0.0.0/8), uptime: 1d20h, pim ip
  Incoming interface: Null, RPF nbr: 0.0.0.0
  Outgoing interface list: (count: 0)

dc3rtg-d2(config-if)# show ip pim group-range
PIM Group-Range Configuration for VRF "default"
Group-range        Mode      RP-address       Shared-tree-only range
232.0.0.0/8        SSM       -                -
231.0.0.0/8        ASM       21.21.0.11       -
231.128.0.0/9      ASM       21.21.0.22       -
231.129.0.0/16     ASM       21.21.0.33       -
231.129.128.0/17   Unknown   -                -
```

Related Documents

Related Topic	Document Title
VDCs	*Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*
CLI commands	*Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*
Configuring VRFs and Policy Based Routing	*Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*

Standards

Standards	Title
No new or modified standards are supported by this feature, and support for existing standards has not been modified by this feature.	—

Feature History for PIM and PIM6

Feature Name	Release	Feature Information
Support for Graceful Restart PIM	6.2(2)	Support for Graceful Restart protocol Independent Multicast (PIM) is a multicast high availability (HA) enhancement that improves the reconvergence of multicast routes (mroutes) after a route processor (RP) switchover. In the event of an RP switchover, this feature uses the PIM-SM Generation ID (GenID) value as a mechanism to trigger adjacent PIM neighbors on an interface to send PIM messages for all (*, G) and (S, G) mroutes that use that interface as an RPF interface, immediately reestablishing those states on the newly active RP.
Support for the **pim register-source** command.	5.2(1)	Support for configuring the IP source address of register messages.
BFD support for PIM (IPv4)	5.0(2)	BFD supported for PIM with IPv4.

Feature Name	Release	Feature Information
vPC	4.1(3)	Cisco NX-OS software for the Nexus 7000 Series devices does not support PIM SSM or BIDR on a vPC. Display vPC statistics with the **show ip pim statistics** command.

CHAPTER 6

Configuring PIM Allow RP

Configuring PIM Allow RP

This chapter describes how to configure the PIM Allow RP feature in IPv4 and IPv6 networks for inter-connecting Protocol Independent Multicast (PIM) Sparse Mode (SM) domains with different rendezvous points (RPs). PIM Allow RP enables the receiving device to use its own RP to create state and build shared trees when an incoming (*, G) Join is processed and a different RP is identified. This allows the receiving device to accept the (*, G) Join from the different RP.

Restrictions for PIM Allow RP

- PIM Allow RP only supports connecting PIM SM domains.
- PIM Allow RP is applicable for downstream traffic only, that is, it is only applicable for building the shared tree.
- PIM Allow RP does not work with Auto-RP or Boot Strap Router (BSR). Only static configuration is supported. However, it does allow the RP used in the consumer network to be different than the one configured statically in the service provider network.
- PIM Allow RP is restricted to use only the route-map.
- PIM Allow RP does not support the IPv6 Multicast prior to Cisco NX-OS Release 8.4(2). IPv6 PIM Allow RP is supported from Cisco NX-OS Release 8.4(2).
- PIM Allow RP does not support the RPM with “Source”. PIM Allow RP Information AboutPIM AllowRP.
- When the Allow-RP configuration is added with a non-existent RPM, all Joins/Prunes get rejected.
- When the Allow-RP configuration is added with an RPM having PERMIT-ALL or DENY-ALL, all Joins/Prunes are either accepted or discarded accordingly.

Information about PIM Allow RP

Rendezvous Points

A rendezvous point (RP) is a role that a device performs when operating in Protocol Independent Multicast (PIM) Sparse Mode (SM). An RP is required only in networks running PIM SM. In the PIM-SM model, only network segments with active receivers that have explicitly requested multicast data will be forwarded the traffic. This method of delivering multicast data is in contrast to PIM Dense Mode (PIM DM). In PIM DM, multicast traffic is initially flooded to all segments of the network. Routers that have no downstream neighbors or directly connected receivers prune back the unwanted traffic. An RP acts as the meeting place for sources and receivers of multicast data. In a PIM-SM network, sources must send their traffic to the RP. This traffic is then forwarded to receivers down a shared distribution tree.

By default, when the first hop device of the receiver learns about the source, it will send a Join message directly to the source, creating a source-based distribution tree from the source to the receiver. This source tree does not include the RP unless the RP is located within the shortest path between the source and receiver. In most cases, the placement of the RP in the network is not a complex decision.

By default, the RP is needed only to start new sessions with sources and receivers. Consequently, the RP experiences little overhead from traffic flow or processing. In PIM version 2, the RP performs less processing than in PIM version 1 because sources must only periodically register with the RP to create state.

PIM Allow RP

There are three types of networks: publisher, consumer, and transport. Many publisher networks can originate content and many consumer networks can be interested in the content. The transport network, owned and operated by a service provider, connects the publisher and the consumer networks.

The consumer and the transport networks are connected as follows: For a specific group range, or all-groups range (similar to a default route), the service provider defines a particular rendezvous point (RP), such as RP-A. Reverse path forwarding of RP-A from a consumer device will cause a (*,G) Join to be sent towards the transport network. For the same group, the service provider may define a different RP, such as RP-B, that is used to build the shared tree within the transport network for G. RP-A and RP-B are typically different RPs and each RP is defined for different group ranges. RFC 4601 dictates that if a device receives a (*, G) Join and the RP that is specified in the (*, G) Join is different than what the receiving device expects (unknown RPs), the incoming (*, G) Join must be ignored.

The PIM Allow RP feature is introduced in Cisco NX-OS Release 8.4(1). This feature enables the receiving device to use its own RP to create state and build shared trees when an incoming (*, G) Join is processed and a different RP is identified. This allows the receiving device to accept the (*, G) Join from the different RP. A route-map is used to control which RP address and/or group addresses the (*,G) join is for. The RP address and the group address in the (*,G) join message is matched against any RP and group addresses specified in the route-map. Support for IPv6 is introduced in Cisco NX-OS Release 8.4(2).

PIM Allow RP is only applicable for downstream traffic, for building the shared tree. It does not work with Auto-RP or BSR. Only static configuration is supported. However, PIM Allow RP does compensate for the embedded RP in the consumer network to be different than the one configured statically in the transport network.

Configuring RPs for PIM-SM

All access lists should be configured prior to beginning the configuration task. For information about how to configure an access list, see the “Configuring IP ACLs” chapter in the *Cisco Nexus 7000 Series NX-OS Security Configuration Guide.*

Procedure

Step 1 Enter the global configuration mode.

config terminal

Step 2 Selects an interface that is connected to hosts on which PIM can be enabled.

interface *type number*

Example: Device(config)# interface gigabitethernet 1/0/0

Step 3 Enable PIM. You must use sparse mode.

ip pim sparse-mode

Example: Device(config-if)# ip pim sparse-mode

Step 4 Enable an interface

no shut

Example: Device(config-if)# no shut

Step 5 Return to global configuration mode.

exit

Example: Device(config-if)# exit

Step 6 Repeat Steps 3 through 6 on every interface that uses IP multicast.

Step 7 Configures a PIM static RP address for a multicast group range. You can specify a route-map policy name that lists the group prefixes to use with the match ip multicast command. This command can also be used in VRF mode.

ip pim rp-address *rp-address* [**group-list***ip-prefix* | **route-map** *policy-name*]

Example: Device(config)# ip pim rp-address 30.2.2.2 group-list 224.0.0.0/4

Step 8 Exit the route map configuration mode.

end

Example: Device(config-route-map)# end

Step 9 (Optional) Display the RPs known in the network and shows how the router learned about each RP.

show ip pim rp[**vrf** | *rp-address*]

Example: Device# show ip pim rp

Step 10 Display the contents of the IP mroute table.

show ip mroute

Example: Device# show ip mroute

Enabling PIM Allow RP

In the following configuration steps, you can configure one of the combinations of RPM at a time —group only, RP only, group RP, group-range only.

Procedure

Step 1 Enter the global configuration mode.

config terminal

Step 2 Enter route-map configuration mode. Note that this configuration method uses the permit keyword.

route-map *map-name* [**permit** | **deny**] [*sequence-number*]

Example: Device(config)# route-map mcast-grp permit 10

Step 3 Match the IP multicast group. Note that you can configure only one combination of RPM at a time —group only, RP only, group RP, group-range only. For example; after you configure this step (group only), you should go to step 9. This is applicable to the below mentioned steps as well (from step-4 to step-8).

match ip multicast group *group-address*

Example: Device(config-route-map)# match ip multicast group 224.0.0.0/4

Step 4 Match the IP multicast group range from/to the specified group address.

match ip multicast group-range{ *group address_start* **to** *group address_end*}

Example: Device(config-route-map)# match ip multicast group-range 230.1.1.1 to 230.1.1.255

Step 5 Match the IP multicast and the RP specified.

match ip multicast rp *rp-address*

Example: Device(config-route-map)# match ip multicast 222.0.0.0/4

Step 6 Match the IP multicast RP address and the RP type specified. ASM is the only supported RP type.

match ip multicast rp *address* **rp-type** *type*

Example: Device(config-route-map)# match ip multicast rp 1.1.1.1/32 rp-type ASM

Step 7 Match the IP multicast group address and the RP address.

match ip multicast group *address* **rp** *address*

Example: Device(config-route-map)# match ip multicast group 230.1.1.1/4 rp 1.1.1.1/32

Step 8 Matches the IP multicast group range from/to the specified address and the RP address.

match ip multicast group-range {*group address_start to group address_end*} **rp** *address*

Example: Device(config-route-map)# match ip multicast group-range 230.1.1.1 to 230.1.1.255 rp 1.1.1.1/32

Step 9 Enable PIM Allow RP; and allow sparse-mode RP addresses. This command is configured at the VRF level also. A route-map is used to control which RP address and/or group addresses the (*,G) join is for. The RP address and the group address in the (*,G) join message is matched against any RP and group addresses specified in the route-map.

ip pim allow-rp *route-map-name*

Example: Device(config-roiute-map)# ip pim allow-rp test-route-map

Step 10 Enable the IPv6 PIM Allow RP.

ipv6 pim allow-rp *route-map-name*

Example: Device(config-roiute-map)# ipv6 pim allow-rp test-route-map

Step 11 Exit the route map configuration mode.

end

Example: Device(config-route-map)# end

Displaying Information About Allow RP Policy

Note: The following commands can be used under VRF mode also.

Procedure

Step 1 Enable privileged EXEC mode.

enable

Example: Device > enable

Step 2 Display the statistics about the current allow RP policy and its counters.

show ip pim policy statistics allow-rp-policy

Example: Device# show ip pim policy statistics allow-rp-policy

Step 3 Display the IPv6 statistics about the current allow RP policy.

show ipv6 pim policy statistics allow-rp-policy

Example: Device# show ipv6 pim policy statistics allow-rp-policy

Step 4 Clears the policy and counters of the allow RP policy.

clear ip pim policy statistics allow-rp-policy

Example: Device# clear ip pim policy statistics allow-rp-policy

Step 5 Clears the policy and counters of the allow RP policy for IPv6.

clear ipv6 pim policy statistics allow-rp-policy

Example: Device# clear ipv6 pim policy statistics allow-rp-policy

Feature Information for PIM Allow RP

This table lists the release history for this feature.

Table 19: Feature Information for PIM Allow RP

Feature Name	Releases	Feature Description
PIM AllowRP (IPv6)	8.4(2)	Support for IPv6 is introduced.
PIM AllowRP	8.4(1)	This feature is introduced. The PIM Allow RP feature enables is processed and a different RP is identified. This process permits the receiving device to accept the (*, G) Join from a different RP.

CHAPTER 7

Configuring IGMP Snooping

This chapter describes how to configure Internet Group Management Protocol (IGMP) snooping on a Cisco NX-OS device.

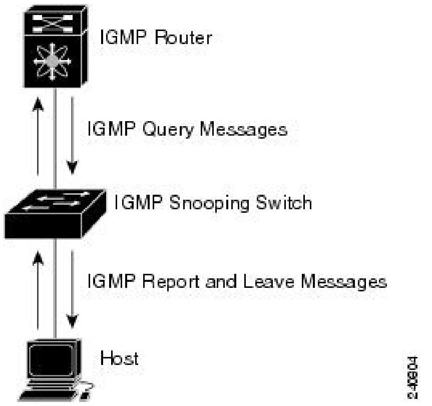
Information About IGMP Snooping

Note We recommend that you do not disable IGMP snooping on the device. If you disable IGMP snooping, you might see reduced multicast performance because of excessive false flooding within the device.

IGMP snooping software examines Layer 2 IP multicast traffic within a VLAN to discover the ports where interested receivers reside. Using the port information, IGMP snooping can reduce bandwidth consumption in a multiaccess LAN environment to avoid flooding the entire VLAN. IGMP snooping tracks which ports are attached to multicast-capable routers to help the routers forward IGMP membership reports. The IGMP snooping software responds to topology change notifications. By default, IGMP snooping is enabled on the device.

This figure shows an IGMP snooping switch that sits between the host and the IGMP router. The IGMP snooping switch snoops the IGMP membership reports and leave messages and forwards them only when necessary to the connected IGMP routers.

Figure 17: IGMP Snooping Switch



The IGMP snooping software operates upon IGMPv1, IGMPv2, and IGMPv3 control plane packets where Layer 3 control plane packets are intercepted and influence the Layer 2 forwarding behavior.

For more information about IGMP, see *Configuring IGMP*.

The Cisco NX-OS IGMP snooping software has the following proprietary features:

- Source filtering that allows forwarding of multicast packets based on destination and source IP.
- Multicast forwarding based on IP addresses rather than MAC addresses.
- Beginning with Cisco Release 5.2(1) for the Nexus 7000 Series devices, multicast forwarding alternately based on the MAC address
- Optimized multicast flooding (OMF) that forwards unknown traffic to routers only and performs no data-driven state creation.

IGMPv1 and IGMPv2

Both IGMPv1 and IGMPv2 support membership report suppression, which means that if two hosts on the same subnet want to receive multicast data for the same group, then the host that receives a member report from the other host suppresses sending its report. Membership report suppression occurs for hosts that share a port.

If no more than one host is attached to each VLAN switch port, you can configure the fast leave feature in IGMPv2. The fast leave feature does not send last member query messages to hosts. As soon as the software receives an IGMP leave message, the software stops forwarding multicast data to that port.

IGMPv1 does not provide an explicit IGMP leave message, so the software must rely on the membership message timeout to indicate that no hosts remain that want to receive multicast data for a particular group.

Note The software ignores the configuration of the last member query interval when you enable the fast leave feature because it does not check for remaining hosts.

IGMPv3

The IGMPv3 snooping implementation on Cisco NX-OS supports full IGMPv3 snooping, which provides constrained flooding based on the (S, G) information in the IGMPv3 reports. This source-based filtering enables the device to constrain multicast traffic to a set of ports based on the source that sends traffic to the multicast group.

By default, the software tracks hosts on each VLAN port. The explicit tracking feature provides a fast leave mechanism. Because every IGMPv3 host sends membership reports, report suppression limits the amount of traffic that the device sends to other multicast-capable routers. When report suppression is enabled, and no IGMPv1 or IGMPv2 hosts requested the same group, the software provides proxy reporting. The proxy feature builds the group state from membership reports from the downstream hosts and generates membership reports in response to queries from upstream queriers.

Even though the IGMPv3 membership reports provide a full accounting of group members on a LAN segment, when the last host leaves, the software sends a membership query. You can configure the parameter last member query interval. If no host responds before the timeout, the software removes the group state.

IGMP Snooping Querier

When PIM is not enabled on an interface because the multicast traffic does not need to be routed, you must configure an IGMP snooping querier to send membership queries. You define the querier in a VLAN that contains multicast sources and receivers but no other active querier.

The querier can be configured to use any IP address in the VLAN.

As a best practice, a unique IP address, one that is not already used by the switch interface or the HSRP VIP, should be configured so as to easily reference the querier. In a vPC configuration too, the querier IP should be unique on the vPC primary and secondary.

Note The IP address for the querier should not be a broadcast IP, multicast IP, or 0(0.0.0.0).

When an IGMP snooping querier is enabled, it sends out periodic IGMP queries that trigger IGMP report messages from hosts that want to receive IP multicast traffic. IGMP snooping listens to these IGMP reports to establish appropriate forwarding.

The IGMP snooping querier performs querier election as described in RFC 2236. A querier election occurs in the following configurations:

- When there are multiple switch queriers configured with the same subnet on the same VLAN on different switches.
- When the configured switch querier is in the same subnet as with other Layer 3 SVI queriers.

Static Multicast MAC Address

Beginning with the Cisco Release 5.2(1) for the Nexus 7000 Series devices, you configure an outgoing interface statically for a multicast MAC address. Also, you can configure the IGMP snooping to use a MAC-based lookup mode.

Previously, the system performs the lookup on a Layer 2 multicast table using the destination IP address rather than the destination MAC address. However, some applications share a single unicast cluster IP and multicast cluster MAC address. The system forwards traffic destined to the unicast cluster IP address by the last-hop router with the shared multicast MAC address. This action can be accomplished by assigning a static multicast MAC address for the destination IP address for the end host or cluster.

The default lookup mode remains IP, but you can configure the lookup type to MAC address-based. You can configure the lookup mode globally or per VLAN:

- If the VDC contains ports from only an M-Series module and the global lookup mode is set to IP, VLANs can be set to either one of the two lookup modes. But, if the global lookup mode is set to a MAC address, the operational lookup mode for all the VLANs changes to MAC-address mode.
- If the VDC contains ports from both an M-Series module and an F-Series module and if you change the lookup mode to a MAC address in any VLAN, the operation lookup mode changes for all of the VLANs to a MAC-address based. With these modules in the chassis, you have the same lookup mode globally and for the VLANs. Similarly, if the global lookup mode is MAC-address based, the operational lookup mode for all VLAN is also MAC-address based.

Note Changing the lookup mode is disruptive. Multicast forwarding is not optimal until all multicast entries are programmed with the new lookup mode. Also, when 32 IP addresses are mapped to a single MAC address, you might see suboptimal forwarding on the device.

IGMP Snooping with VDCs and VRFs

A virtual device context (VDC) is a logical representation of a set of system resources. Within each VDC, you can define multiple virtual routing and forwarding (VRF) instances. One IGMP process can run per VDC. The IGMP process supports all VRFs in that VDC and performs the function of IGMP snooping within that VDC.

You can use the *show* commands with a VRF argument to provide a context for the information displayed. The default VRF is used if no VRF argument is supplied.

For information about configuring VDCs, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*.

For information about configuring VRFs, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

IGMP Snooping across VPLS Domains

Beginning with Cisco Release 6.2(2) for the Nexus 7000 Series devices, IGMP snooping can be configured across Virtual Private LAN Service (VPLS) domains. The IGMP Snooping across VPLS Domains feature enables snooping of the IGMP packets on the pseudowire and on the Layer 2 side of the network for optimal delivery of the multicast packets.

A pseudowire is a point-to-point connection between pairs of Provider Edge (PE) devices. A pseudowire emulates services like Ethernet over an underlying core multiprotocol label switching (MPLS) network through encapsulation into a common MPLS format. A pseudowire allows carriers to converge their services to an MPLS network by encapsulating services into a common MPLS format.

By snooping IGMP packets received on a link, the device sends multicast packets only to interested end points. Once an IGMP packet going over the Layer 2 link is snooped, it is passed to the control plane. The control plane will add the link on which it was received to the multicast group. The IGMP packets coming on the pseudowire are also snooped and sent to the control plane. The control plane then adds the pseudowire to the multicast group. When a multicast packet is received, it will be sent only to the multicast group instead of flooding the VLAN.

Prerequisites for IGMP Snooping

IGMP snooping has the following prerequisites:

- You are logged onto the device.
- You are in the correct virtual device context (VDC). A VDC is a logical representation of a set of system resources. You can use the **switchto vdc** command with a VDC number.
- For global commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.

Guidelines and Limitations for IGMP Snooping

IGMP snooping has the following guidelines and limitations:

- You must disable IGMP optimized multicast flooding (OMF) for IPv6 multicast networks that require multicast forwarding over a layer 2 network.
- You must disable IGMP optimized multicast forwarding on VLANs that require forwarding of IPv6 packets.
- When a vPC peer-link runs in a F2 module, IGMP querier election does not happen. Hence do not configure vPC peer-link in a F2 module.
- If you are configuring vPC peers, the differences in the IGMP snooping configuration options between the two devices have the following results:
    - If IGMP snooping is enabled on one device but not on the other, the device on which snooping is disabled floods all multicast traffic.
    - A difference in multicast router or static group configuration can cause traffic loss.
    - The fast leave, explicit tracking, and report suppression options can differ if they are used for forwarding traffic.
    - If a query parameter is different between the devices, one device expires the multicast state faster while the other device continues to forward. This difference results in either traffic loss or forwarding for an extended period.
    - If an IGMP snooping querier is configured on both devices, only one of them will be active because an IGMP snooping querier shuts down if a query is seen in the traffic.
- You must enable ip igmp snooping group-timeout when you use ip igmp snooping proxy general-queries. We recommend to set it to "never." If this is not done you might have multicast packet loss.

- Network applications that use unicast destination IP addresses with multicast destination MAC addresses might require the configuration of IGMP snooping to use MAC-based forwarding lookups on the switch. If the destination MAC address used for this kind of applications is a non-IP multicast MAC address, use the **mac address-table multicast** command to statically configure the port membership. If the destination MAC address is in the IP multicast range, 0100.5E00.0000 to 0100.5E7F.FFFF, use static IGMP snooping membership entries for the corresponding Layer 3 IP multicast address to configure the port membership. For example, if the application uses destination MAC address 0100.5E01.0101, configure a static IGMP snooping membership entry for an IP multicast address that maps to that MAC address. An example of this is **ip igmp snooping static-group 239.1.1.1**.

Default Settings for IGMP Snooping

This table lists the default settings for IGMP snooping parameters.

Parameters	Default
IGMP snooping	Enabled
Explicit tracking	Enabled
Fast leave	Disabled
Last member query interval	1 second
Snooping querier	Disabled
Report suppression	Enabled
Link-local groups suppression	Enabled
IGMPv3 report suppression for the entire device	Disabled
IGMPv3 report suppression per VLAN	Enabled

Configuring IGMP Snooping Parameters

Note If you are familiar with the Cisco IOS CLI, be aware that the Cisco NX-OS commands for this feature might differ from the Cisco IOS commands that you would use.

Note You must enable IGMP snooping globally before any other commands take effect.

Configuring Global IGMP Snooping Parameters

To affect the operation of the IGMP snooping process globally, you can configure the optional IGMP snooping parameters described in the following table:

Parameter	Description
IGMP snooping	Enables IGMP snooping on the active VDC. The default is enabled. **Note** If the global setting is disabled, all VLANs are treated as disabled, whether they are enabled or not.
Event history	Configures the size of the IGMP snooping history buffers. The default is small.
Group timeout	Configures the group membership timeout for all VLANs on the device.
Link-local groups suppression	Configures link-local groups suppression on the device. The default is enabled.
Optimise-multicast-flood	Configures Optimized Multicast Flood (OMF) on all VLANs on the device. The default is enabled.
Proxy	Configures IGMP snooping proxy for the device. The default is 5 seconds.
Report suppression	Limits the membership report traffic sent to multicast-capable routers on the device. When you disable report suppression, all IGMP reports are sent as is to multicast-capable routers. The default is enabled.
IGMPv3 report suppression	Configures IGMPv3 report suppression and proxy reporting on the device. The default is disabled.

Notes for IGMP Snooping Parameters

The following are additional notes about some of the IGMP snooping parameters.

- IGMP Snooping Proxy parameter

  To decrease the burden placed on the snooping switch during each IGMP general query (GQ) interval, Cisco NX-OS provides a way to decouple the periodic general query behavior of the IGMP snooping switch from the query interval configured on the multicast routers.

  Beginning with Cisco NX-OS release 5.2(1), a configuration option became available to enable the Cisco Nexus 7000 switch to consume IGMP general queries from the multicast router, rather than flooding the general queries to all the switchports.

  When receiving a general query, the switch produces proxy reports for all currently active groups and distributes the proxy reports over the period specified by the MRT that is specified in the router query. At the same time, independent of the periodic general query activity of the multicast router, the switch sends an IGMP general query on each port in the VLAN in a round-robin fashion. It cycles through all the interfaces in the VLAN at the rate given by the following formula.

Rate = {number of interfaces in VLAN} * {configured MRT} * {number of VLANS}

When running queries in this mode, the default MRT value is 5,000 milliseconds (5 seconds), which means that in a switch that has 500 switchports in a VLAN, it would take 2,500 seconds (40 minutes) to cycle through all the interfaces in the system. This is also true when the Cisco Nexus 7000 switch itself is the querier.

This behavior ensures that only one host responds to a general query at a given time and it keeps the simultaneous reporting rate below the packet-per-second IGMP capability of the switch (approximately 3,000 to 4,000 pps).

Note When using this option, you must change the **ip igmp snooping group-timeout** parameter to a high value or to never time out.

The **ip igmp snooping proxy general-queries**[*mrt*] command causes the snooping function to proxy reply to general queries from the multicast router, while also sending round-robin general queries on each switchport with the specified MRT value (the default MRT value is 5 seconds).

- IGMP Snooping Group-timeout parameter

  Configuring the group-timeout parameter disables the behavior of expiring membership based on three missed general queries. The group membership remains on a given switchport until the switch receives an explicit IGMP leave on that port.

  The **ip igmp snooping group-timeout** {*timeout*|*never*} command modifies or disables the behavior of an expiring IGMP snooping group membership after three missed general queries.

Procedure

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td><b>Step 1</b></td><td><b>config t</b>
<b>Example:</b>
<code>switch# config t</code>
<code>switch(config)#</code></td><td>Enters global configuration mode.</td></tr>
<tr><td><b>Step 2</b></td><td><table><tr><th>Option</th><th>Description</th></tr><tr><td><b><code>ip igmp snooping</code></b>
<code>switch(config)# ip igmp snooping</code></td><td>Enables IGMP snooping for the device. The default is enabled.</td></tr></table></td><td>The following commands can be used to configure the IGMP snooping.</td></tr>
</table>

	Command or Action	Purpose

Option	Description
	Note If the global setting is disabled with the no form of this command, IGMP snooping on all VLANs is disabled, whether IGMP snooping is enabled on a VLAN or not. If you disable IGMP snooping, Layer 2 multicast frames flood to all modules. **Note** IGMP snooping can be configured across Virtual Private LAN Service (VPLS) domains.
`ip igmp snooping event-history {igmp-snoop-internal | mfdm | mfdm-sum | rib | vlan | vlan-events | vpc} size`	Configures the size of the event history buffer. The default is **small.**

	Command or Action	Purpose

Option	Description

{**disabled** | **large** | **medium** | **small**}
`switch(config)# ip igmp snooping event-history igmp-snoop-internal size large`	

ip igmp snooping group-timeout{*minutes*|*never*} `switch(config)# ip igmp snooping group-timeout never`	Configures the group membership timeout value for all VLANs on the device.
ip igmp snooping link-local-groups-suppression `switch(config)# ip igmp snooping link-local-groups-suppression`	Configures link-local groups suppression for the entire device. The default is enabled.
ip igmp snooping optimise-multicast-flood `switch(config)# ip igmp snooping optimise-multicast-flood`	Optimizes OMF on all VLANs on the device. The default is enabled.
ip igmp snooping proxy general-queries [**mrt** *seconds*] `switch(config)# ip igmp snooping proxy general-queries`	Configures IGMP snooping proxy for the device. The default is 5 seconds.
ip igmp snooping v3-report-suppression `switch(config)# ip igmp snooping v3-report-suppression`	Limits the membership report traffic sent to multicast-capable routers. When you disable report suppression, all IGMP reports are sent as is to multicast-capable routers. The default is enabled.
ip igmp snooping report-suppression `switch(config)# ip igmp snooping report-suppression`	Configures IGMPv3 report suppression and proxy reporting. The default is disabled.

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td></td><td colspan="2"><table><tr><th>Option</th><th>Description</th></tr><tr><td><b>ip igmp snooping max-gq-miss</b> <i>count</i>
switch(config)# ip igmp snooping max-gq-miss 5</td><td>Configures the maximum number of general query misses permitted. The range is 3 to 5 queries. The default is 3 queries.</td></tr></table></td></tr>
<tr><td><b>Step 3</b></td><td><b>copy running-config startup-config</b>
<b>Example:</b>
switch(config)# copy running-config startup-config</td><td>(Optional) Saves configuration changes.</td></tr>
</table>

Configuring IGMP Snooping Parameters per VLAN

To affect the operation of the IGMP snooping process per VLAN, you can configure the optional IGMP snooping parameters described in this table.

Parameter	Description
IGMP snooping	Enables IGMP snooping on a per-VLAN basis. The default is enabled. **Note** If the global setting is disabled, all VLANs are treated as disabled, whether they are enabled or not.
Explicit tracking	Tracks IGMPv3 membership reports from individual hosts for each port on a per-VLAN basis. The default is enabled.
Fast leave	Enables the software to remove the group state when it receives an IGMP leave report without sending an IGMP query message. This parameter is used for IGMPv2 hosts when no more than one host is present on each VLAN port. The default is disabled.
Group timeout	Configures the group membership timeout for the specified VLANs.
Last member query interval	Sets the interval that the software waits after sending an IGMP query to verify that no hosts that want to receive a particular multicast group remain on a network segment. If no hosts respond before the last member query interval expires, the software removes the group from the associated VLAN port. Values range from 1 to 25 seconds. The default is 1 second.

Parameter	Description
Optimise-multicast-flood	Configures Optimized Multicast Flood (OMF) on specified VLANs. The default is enabled.
Proxy	Configures IGMP snooping proxy for the specified VLANs. The default is 5 seconds.
Snooping querier	Configures a snooping querier on an interface when you do not enable PIM because multicast traffic does not need to be routed. You can also configure the following values for the snooping querier: • timeout—Timeout value for IGMPv2 • interval—Time between query transmissions • maximum response time—MRT for query messages • startup count—Number of queries sent at startup • startup interval—Interval between queries at startup
Robustness variable	Configures the robustness value for the specified VLANs.
Report suppression	Limits the membership report traffic sent to multicast-capable routers on a per-VLAN basis. When you disable report suppression, all IGMP reports are sent as is to multicast-capable routers. The default is enabled.
Multicast router	Configures a static connection to a multicast router. The interface to the router must be in the selected VLAN.
Static group	Configures a Layer 2 port of a VLAN as a static member of a multicast group.
Link-local groups suppression	Configures link-local groups suppression on a per-VLAN basis. The default is enabled.
IGMPv3 report suppression	Configures IGMPv3 report suppression and proxy reporting on a per-VLAN basis. The default is enabled per VLAN.
Version	Configures the IGMP version number for the specified VLANs. **Note** You must configure access-group (policy filter), for this command to function correctly.

Note Beginning with Cisco Release 5.1(1), step 3 in the following procedure changed from **vlan** to **vlan configuration** *vlan-id*. You configure the IP IGMP snooping parameters that you want by using this configuration mode; however, the configurations apply only after you specifically create the specified VLAN. See the *Cisco Nexus 7000 Series NX-OS Layer 2 Switching Configuration Guide* for information on creating VLANs.

Procedure

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td><b>Step 1</b></td><td><b>configure terminal</b>
<b>Example:</b>
<code>switch# configure terminal</code>
<code>switch(config)#</code></td><td>Enters global configuration mode.</td></tr>
<tr><td><b>Step 2</b></td><td><b>ip igmp snooping</b>
<b>Example:</b>
<code>switch(config)# ip igmp snooping</code></td><td>Enables IGMP snooping for the current VDC. The default is enabled.
<b>Note</b> If the global setting is disabled with the <b>no</b> form of this command, IGMP snooping on all VLANs is disabled, whether IGMP snooping is enabled on a VLAN or not. If you disable IGMP snooping, Layer 2 multicast frames flood to all modules.</td></tr>
<tr><td><b>Step 3</b></td><td><table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td><b>vlan</b> <i>vlan-id</i>
<code>switch(config)# vlan 2</code>
<code>switch(config-vlan)#</code></td><td>Enters VLAN configuration mode.</td></tr>
<tr><td><b>vlan configuration</b><i>vlan-id</i>
<code>switch(config)# vlan configuration 2</code>
<code>switch(config-vlan-config)#</code></td><td>Beginning with Cisco Release 5.1(1), use this command to configure the IGMP snooping parameters you want for the VLAN. These configurations do not apply until you</td></tr>
</table></td><td>Depending on your release of Cisco NX-OS, use one of the commands in the table.</td></tr>
</table>

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td></td><td><table><tr><th>Option</th><th>Description</th></tr><tr><td></td><td>create the specified VLAN.</td></tr></table></td><td></td></tr>
<tr><td>Step 4</td><td><table><tr><th>Option</th><th>Description</th></tr><tr><td><b>ip igmp snooping</b>
<code>switch(config-vlan-config)# ip igmp snooping</code></td><td>Enables IGMP snooping for the current VLAN. The default is enabled.</td></tr><tr><td><b>ip igmp snooping explicit-tracking</b>
<code>switch(config-vlan-config)# ip igmp snooping explicit-tracking</code></td><td>Tracks IGMPv3 membership reports from individual hosts for each port on a per-VLAN basis. The default is enabled on all VLANs.</td></tr><tr><td><b>ip igmp snooping fast-leave</b>
<code>switch(config-vlan-config)# ip igmp snooping fast-leave</code></td><td>Supports IGMPv2 hosts that cannot be explicitly tracked</td></tr></table></td><td>These commands configure IGMP snooping parameters.</td></tr>
</table>

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td></td><td>
<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td></td><td>because of the host report suppression mechanism of the IGMPv2 protocol. When you enable fast leave, the IGMP software assumes that no more than one host is present on each VLAN port. The default is disabled for all VLANs.</td></tr>
<tr><td><b>ip igmp snooping group-timeout</b> {<i>minutes</i> | <b>never</b>}
<code>switch(config-vlan-config)# ip igmp snooping group-timeout never</code></td><td>Configures the group membership timeout for the</td></tr>
</table>
</td><td></td></tr>
</table>

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td></td><td>
<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td></td><td>specified VLANs</td></tr>
<tr><td><b>ip igmp snooping last-member-query-interval</b> <i>seconds</i>
switch(config-vlan-config)# ip igmp snooping last-member-query-interval 3</td><td>Removes the group from the associated VLAN port if no hosts respond to an IGMP query message before the last member query interval expires. Values range from 1 to 25 seconds. The default is 1 second.</td></tr>
<tr><td><b>ip igmp snooping optimise-multicast-flood</b>
switch(config-vlan-config)# ip igmp snooping optimise-multicast-flood</td><td>Optimizes OMF on selected VLANs. The default is enabled.</td></tr>
</table>
</td><td></td></tr>
</table>

	Command or Action	Purpose

Option	Description
ip igmp snooping proxy general-queries mrt *seconds* `switch(config-vlan-config)# ip igmp snooping proxy general-queries`	Configures an IGMP snooping proxy for specified VLANs. The default is 5 seconds.
ip igmp snooping querier *ip-address* `switch(config-vlan-config)# ip igmp snooping querier 172.20.52.106`	Configures a snooping querier when you do not enable PIM because multicast traffic does not need to be routed. The IP address is used as the source in messages.
ip igmp snooping querier-timeout *seconds* `switch(config-vlan-config)# ip igmp snooping querier-timeout 300`	Configures a snooping querier timeout

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td></td><td>
<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td></td><td>value for IGMPv2 when you do not enable PIM because multicast traffic does not need to be routed The default is 255 seconds</td></tr>
<tr><td><b>ip igmp snooping query-interval</b> <i>seconds</i>
switch(config-vlan-config)# ip igmp snooping query-interval 120</td><td>Configures a snooping query interval when you do not enable PIM because multicast traffic does not need to be routed The default value is</td></tr>
</table>
</td><td></td></tr>
</table>

	Command or Action	Purpose

Option	Description
	125 seconds.
ip igmp snooping query-max-response-time *seconds* `switch(config-vlan-config)# ip igmp snooping query-max-response-time 12`	Configures a snooping MRT for query messages when you do not enable PIM because multicast traffic does not need to be routed. The default value is 10 seconds.
ip igmp snooping startup-query-count *value* `switch(config-vlan-config)# ip igmp snooping startup-query-count 5`	Configures snooping for a number of queries sent at startup when you do not enable PIM

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td></td><td>
<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td></td><td>because multicast traffic does not need to be routed.</td></tr>
<tr><td><b>ip igmp snooping startup-query-interval</b> <i>seconds</i>
switch(config-vlan-config)# ip igmp snooping startup-query-interval 15000</td><td>Configures a snooping query interval at startup when you do not enable PIM because multicast traffic does not need to be routed.</td></tr>
<tr><td><b>ip igmp snooping robustness-variable</b> <i>value</i>
switch(config-vlan-config)# ip igmp snooping robustness-variable 5</td><td>Configures the robustness value for the specified VLANs. The default value is 2.</td></tr>
<tr><td><b>ip igmp snooping report-suppression</b></td><td>Limits the</td></tr>
</table>
</td><td></td></tr>
</table>

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td></td><td>
<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td><code>switch(config-vlan-config)# ip igmp snooping report-suppression</code></td><td>membership report traffic sent to multicast-capable routers. When you disable report suppression, all IGMP reports are sent as is to multicast-capable routers. The default is enabled.</td></tr>
<tr><td><b>ip igmp snooping mrouter interface</b> <i>interface</i>
<code>switch(config-vlan-config)# ip igmp snooping mrouter interface ethernet 2/1</code></td><td>Configures a static connection to a multicast router. The interface to the router must be in the selected VLAN. You can specify</td></tr>
</table>
</td><td></td></tr>
</table>

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td></td><td>
<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td></td><td>the interface by the type and the number, such as <b>ethernet</b> <i>slot/port</i></td></tr>
<tr><td><b>ip igmp snooping static-group</b> [<i>group-ip-addr</i>]<b>source</b> [<i>source-ip-addr</i>] <b>interface</b> <i>interface</i>
<code>switch(config-vlan-config)# ip igmp snooping static-group 230.0.0.1 interface ethernet 2/1</code></td><td>Configures a Layer 2 port of a VLAN as a static member of a multicast group. You can specify the interface by the type and the number, such as <b>ethernet</b> <i>slot/port</i></td></tr>
<tr><td><b>ip igmp snooping link-local-groups-suppression</b>
<code>switch(config-vlan-config)# ip igmp snooping link-local-groups-suppression</code></td><td>Configures link-local groups suppression for the</td></tr>
</table>
</td><td></td></tr>
</table>

	Command or Action	Purpose

Option	Description
	specified VLANs. The default is enabled.
ip igmp snooping v3-report-suppression `switch(config-vlan-config)# ip igmp snooping v3-report-suppression`	Configures IGMPv3 report suppression and proxy reporting for the specified VLANs. The default is enabled per VLAN.
ip igmp snooping version *value* `switch(config-vlan-config)# ip igmp snooping version 2`	Configures the IGMP version number for the specified VLANs. **Note** You must configure access-group (policy filter), for this command to function correctly.

	Command or Action	Purpose
Step 5	**copy running-config startup-config** **Example:**	(Optional) Saves configuration changes.

	Command or Action	Purpose
	`switch(config)# copy running-config startup-config`	

Changing the Lookup Mode

Beginning with Cisco Release 5.2(1) for the Nexus 7000 Series chassis, you can configure the lookup mode to be based on the MAC address either globally or per VLAN.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**layer-2 multicast lookup mac** **Example:** `switch(config)# layer-2 multicast lookup mac`	Globally changes the lookup mode to be based on the MAC address. To return to the default IP lookup mode, use the **no** form of this command. **Note** After **layer-2 multicast lookup mac** is configured, the Cisco Nexus 7000 device still floods unicast traffic with multicast MAC address under the following conditions: • Both ingress and egress ports are either on M1 or M2 module. • Both ingress and egress ports are layer 2 ports (e.g. either an access port or a trunk port) in two different VLANs. Cisco Nexus 7000 device provides routing between the two VLANs. • The destination IP address is a NLB multicast/IGMP host. In other words, the destination IP is unicast and the destination MAC address starts with 0100.5E.
Step 3	**vlan** *vlan-id* **Example:**	Changes the lookup mode to be based on the MAC address for the specified VLANs. To return to the default IP lookup mode for these VLANs, use the **no** form of this command.

	Command or Action	Purpose

	`switch(config)# vlan 5`
	`switch(config-vlan)#`
	`layer-2 multicast lookup mac`
	`switch(config-vlan)# layer-2 multicast lookup mac`
	`switch(config-vlan)#`	

Step 4	**exit**
	Example:
	`switch(config)# exit`
	`switch#`
		Exits configuration and/or VLAN configuration mode.

Step 5	**show ip igmp snooping lookup-mode vlan** [*vlan-id*] **Example:** `switch# show ip igmp snooping lookup-mode`	(Optional) Displays the IGMP snooping lookup mode.
Step 6	**copy running-config startup-config** **Example:** `switch# copy running-config startup-config`	(Optional) Copies the running configuration to the startup configuration.

Configuring a Static Multicast MAC Address

Beginning with Cisco Release 5.2(1) for the Nexus 7000 Series chassis, you can configure an outgoing interface statically for a multicast MAC address.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**mac address-table multicast** *multicast-mac-addr* **vlan** *vlan-id* **interface** *slot/port* **Example:** `switch(config)# mac address-table`	Configures the specified outgoing interface statically for a multicast MAC address.

	Command or Action	Purpose

	`multicast 01:00:5f:00:00:00 vlan 5`
	`interface ethernet 2/5`	

Step 3	**exit**
	Example:
	`switch(config)# exit`
	`switch#`
		Exits configuration and/or VLAN configuration mode.

Step 4	**show ip igmp snooping mac-oif** [**detail** | **vlan** *vlan-id* [**detail**]] **Example:** `switch# show feature-set`	(Optional) Displays the IGMP snooping static MAC addresses.

Step 5	**copy running-config startup-config**
	Example:
	`switch# copy running-config`
	`startup-config`
		(Optional) Copies the running configuration to the startup configuration.

Verifying IGMP Snooping Configuration

To display the IGMP configuration information, perform one of the following tasks:

Command or Action	Purpose
show ip igmp snooping [**vlan** *vlan-id*]	Displays the IGMP snooping configuration by VLAN.
show ip igmp snooping groups [*source* [*group*] | *group* | [*source*] | [**vlan** *vlan-id*] [**detail**]	Displays IGMP snooping information about groups by VLAN.
show ip igmp snooping querier [**vlan** *vlan-id*]	Displays IGMP snooping queriers by VLAN.
show ip igmp snooping mroute [**vlan** *vlan-id*]	Displays multicast router ports by VLAN.
show ip igmp snooping explicit-tracking [**vlan** *vlan-id*]	Displays IGMP snooping explicit tracking information by VLAN.
show ip igmp snooping lookup-mode [**vlan** *vlan-id*]	Displays the IGMP snooping lookup mode.
show ip igmp snooping mac-oif [**detail** | **vlan** *vlan-id*[**detail**]]	Displays IGMP snooping static MAC addresses.
show ip igmp snooping pw vlan brief	Displays VLANs, which have pseudowire interfaces that are operationally up.

Displaying IGMP Snooping Statistics

Use the **show ip igmp snooping statistics vlan** command to display IGMP snooping statistics. You can see the virtual port channel (vPC) statistics in this output.

Use the **clear ip igmp snooping statistics vlan** command to clear IGMP snooping statistics.

For detailed information about using these commands, see the *Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*.

Configuration Example for IGMP Snooping

This example shows how to configure the IGMP snooping parameters:

```
switch# config t
switch# ip igmp snooping
switch# vlan 2
switch# ip igmp snooping
switch# ip igmp snooping explicit-tracking
switch# ip igmp snooping fast-leave
switch# ip igmp snooping last-member-query-interval 3
switch# ip igmp snooping querier 172.20.52.106
switch# ip igmp snooping report-suppression
switch# ip igmp snooping mrouter interface ethernet 2/1
switch# ip igmp snooping static-group 230.0.0.1 interface ethernet 2/1
switch# ip igmp snooping link-local-groups-suppression
switch# ip igmp snooping v3-report-suppression
```

This example shows how to configure the IGMP snooping parameters beginning with Cisco Release 5.1(1):

```
switch# config t
switch# ip igmp snooping
switch# vlan configuration 2
switch# ip igmp snooping
switch# ip igmp snooping explicit-tracking
switch# ip igmp snooping fast-leave
switch# ip igmp snooping last-member-query-interval 3
switch# ip igmp snooping querier 172.20.52.106
switch# ip igmp snooping report-suppression
switch# ip igmp snooping mrouter interface ethernet 2/1
switch# ip igmp snooping static-group 230.0.0.1 interface ethernet 2/1
switch# ip igmp snooping link-local-groups-suppression
switch# ip igmp snooping v3-report-suppression
```

The following example shows how to configure IGMP Snooping across VPLS Domains:

```
switch# configure terminal
switch(config)# ip igmp snooping
switch(config)# ip igmp snooping event-history igmp-snoop-internal size large
switch(config)# ip igmp snooping group-timeout never
switch(config)# ip igmp snooping link-local-groups-suppression
switch(config)# ip igmp snooping optimise-multicast-flood
switch(config)# ip igmp snooping proxy general-queries
switch(config)# ip igmp snooping report-suppression
switch(config)# ip igmp snooping v3-report-suppression
```

These configurations do not apply until you specifically create the VLAN. See the *Cisco Nexus 7000 Series NX-OS Layer 2 Switching Configuration Guide* for information on creating VLANs.

Related Documents

Related Topic	Document Title
VDCs	*Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*
CLI commands	*Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*

Standards

Standards	Title
No new or modified standards are supported by this feature, and support for existing standards has not been modified by this feature.	—

Feature History for IGMP Snooping in CLI

Feature Name	Releases	Feature Information
ip igmp snooping max-gq-miss *count*	6.2(2)	Command added to allow you to configure the maximum number of general query misses permitted.
IGMP Snooping across VPLS domains	6.2(2)	The IGMP Snooping across VPLS Domains feature enables snooping of the IGMP packets on the pseudowire as well as on the Layer 2 side of the network for optimal delivery of the multicast packets. The following command was introduced: **show ip igmp snooping pw vlan brief**

Feature Name	Releases	Feature Information
Configuring lookup mode to MAC and assigning a static MAC address	5.2(1)	You can configure IGMP snooping to use the forwarding lookup mode as MAC-based, as well as assign a static MAC address.
vlan configuration *vlan-id*	5.1(1)	Command added to allow you to configure a VLAN before you actually create the VLAN.
vPC	4.1(3)	List of guidelines and limitations that apply to a vPC. Display vPC statistics with the show ip igmp snooping statistics vlan command. The following sections provide information about this feature: • *Guidelines and Limitations for IGMP Snooping* • *Displaying IGMP Snooping Statistics*

CHAPTER 8

Configuring MSDP

This chapter describes how to configure Multicast Source Discovery Protocol (MSDP) on a Cisco NX-OS device.

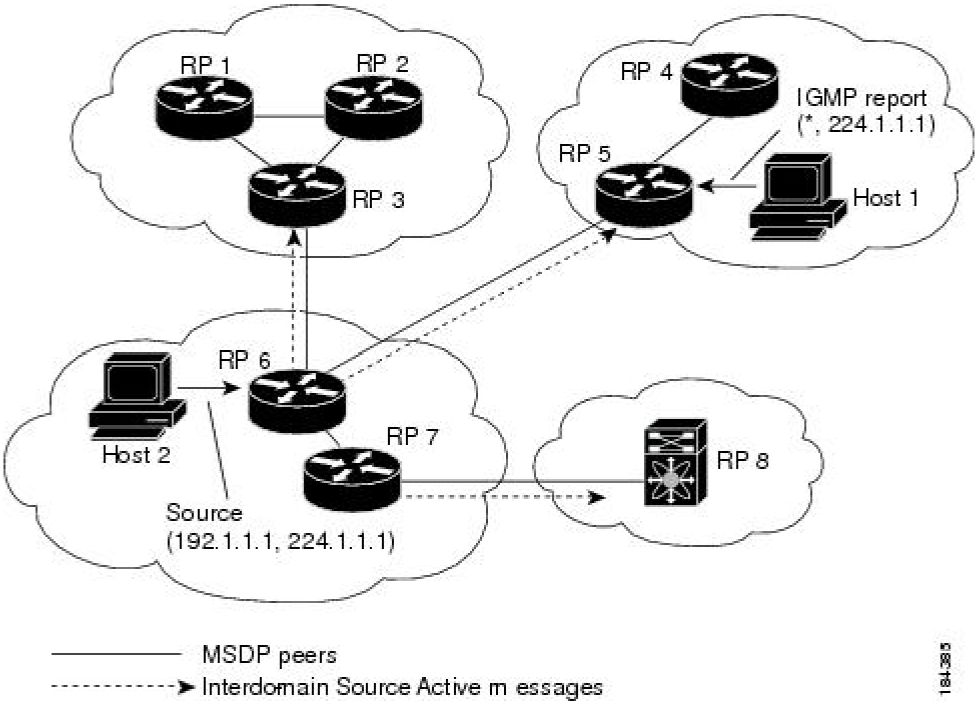
Information About MSDP

You can use the Multicast Source Discovery Protocol (MSDP) to exchange multicast source information between multiple BGP-enabled Protocol Independent Multicast (PIM) sparse-mode domains. In addition, MSDP can be used to create an Anycast-RP configuration to provide RP redundancy and load sharing. For information about PIM, see *Configuring PIM and PIM6*. For information about BGP, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

When a receiver joins a group that is transmitted by a source in another domain, the rendezvous point (RP) sends PIM join messages in the direction of the source to build a shortest path tree. The designated router (DR) sends packets on the source-tree within the source domain, which may travel through the RP in the source domain and along the branches of the source-tree to other domains. In domains where there are receivers, RPs in those domains can be on the source-tree. The peering relationship is conducted over a TCP connection.

The following figure shows four PIM domains. The connected RPs (routers) are called MSDP peers because they are exchanging active source information with each other. Each MSDP peer advertises its own set of multicast source information to the other peers. Source Host 2 sends the multicast data to group 224.1.1.1. On RP 6, the MSDP process learns about the source through PIM register messages and generates Source-Active (SA) messages to its MSDP peers that contain information about the sources in its domain. When RP 3 and RP 5 receive the SA messages, they forward them to their MSDP peers. When RP 5 receives the request from Host 1 for the multicast data on group 224.1.1.1, it builds a shortest path tree to the source by sending a PIM join message in the direction of Host 2 at 192.1.1.1.

Figure 18: MSDP Peering Between RPs in Different PIM Domains



When you configure MSDP peering between each RP, you create a full mesh. Full MSDP meshing is typically done within an autonomous system, as shown between RPs 1, 2, and 3, but not across autonomous systems. You use BGP to do loop suppression and MSDP peer-RPF to suppress looping SA messages.

Note You do not need to configure BGP in order to use Anycast-RP (a set of RPs that can perform load balancing and failover) within a PIM domain.

Note You can use PIM Anycast (RFC 4610) to provide the Anycast-RP function instead of MSDP. For information about PIM, see *Configuring PIM and PIM6*.

For detailed information about MSDP, see RFC 3618.

SA Messages and Caching

MSDP peers exchange Source-Active (SA) messages to propagate information about active sources. SA messages contain the following information:

- Source address of the data source
- Group address that the data source uses
- IP address of the RP or the configured originator ID

When a PIM register message advertises a new source, the MSDP process reencapsulates the message in an SA message that is immediately forwarded to all MSDP peers.

The SA cache holds the information for all sources learned through SA messages. Caching reduces the join latency for new receivers of a group because the information for all known groups can be found in the cache. You can limit the number of cached source entries by configuring the SA limit peer parameter. You can limit the number of cached source entries for a specific group prefix by configuring the group limit global parameter. The SA cache is enabled by default and cannot be disabled.

The MSDP software sends SA messages for each group in the SA cache every 60 seconds or at the configured SA interval global parameter. An entry in the SA cache is removed if an SA message for that source and group is not received within SA interval plus 3 seconds.

MSDP Peer-RPF Forwarding

MSDP peers forward the SA messages that they receive away from the originating RP. This action is called peer-RPF flooding. The router examines the BGP or MBGP routing table to determine which peer is the next hop in the direction of the originating RP of the SA message. This peer is called a reverse path forwarding (RPF) peer.

If the MSDP peer receives the same SA message from a non-RPF peer in the direction of the originating RP, it drops the message. Otherwise, it forwards the message to all its MSDP peers.

MSDP Mesh Groups

You can use MSDP mesh groups to reduce the number of SA messages that are generated by peer-RPF flooding. By configuring a peering relationship between all the routers in a mesh and then configuring a mesh group of these routers, the SA messages that originate at a peer are sent by that peer to all other peers. SA messages received by peers in the mesh are not forwarded.

A router can participate in multiple mesh groups. By default, no mesh groups are configured.

Virtualization Support

A virtual device context (VDC) is a logical representation of a set of system resources. Within each VDC, you can define multiple virtual routing and forwarding (VRF) instances. The MSDP configuration applies to the VRF selected within the current VDC.

You can use the **show** commands with a VRF argument to provide a context for the information displayed. The default VRF is used if no VRF argument is supplied.

For information about configuring VDCs, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*.

For information about configuring VRFs, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

Prerequisites for MSDP

MSDP has the following prerequisites:

- You are logged onto the device.

- You are in the correct virtual device context (VDC). A VDC is a logical representation of a set of system resources. You can use the **switchto vdc** command with a VDC number.
- For global commands, you are in the correct virtual routing and forwarding (VRF) mode. The default configuration mode shown in the examples in this chapter applies to the default VRF.
- You configured PIM for the networks where you want to configure MSDP.

Default Settings for MSDP

This table lists the default settings for the MSDP parameters.

Table 20: Default MSDP Parameters

Parameters	Default
Description	Peer has no description
Administrative shutdown	Peer is enabled when it is defined
MD5 password	No MD5 password is enabled
SA policy IN	All SA messages are received
SA policy OUT	All registered sources are sent in SA messages
SA limit	No limit is defined
Originator interface name	RP address of the local system
Group limit	No group limit is defined
SA interval	60 seconds

Configuring MSDP

You can establish MSDP peering by configuring the MSDP peers within each PIM domain.

1. Select the routers to act as MSDP peers.
2. Enable the MSDP feature.
3. Configure the MSDP peers for each router identified in Step 1.
4. Configure the optional MSDP peer parameters for each MSDP peer.
5. Configure the optional global parameters for each MSDP peer.
6. Configure the optional mesh groups for each MSDP peer.

Note: The MSDP commands that you enter before you enable MSDP are cached and then run when MSDP is enabled. Use the **ip msdp peer** or **ip msdp originator-id** command to enable MSDP.

Note: If you are familiar with the Cisco IOS CLI, be aware that the Cisco NX-OS commands for this feature might differ from the Cisco IOS commands that you would use.

Enabling the MSDP Feature

Before you begin

Before you can access the MSDP commands, you must enable the MSDP feature.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**feature msdp** **Example:** `switch# feature msdp`	Enables the MSDP feature so that you can enter MSDP commands. By default, the MSDP feature is disabled.
Step 3	**show running-configuration | grep** *feature* **Example:** `switch# show running-configuration | grep feature`	(Optional) Displays feature commands that you specified.
Step 4	**copy running-config startup-config** *feature* **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring MSDP Peers

You can configure an MSDP peer when you configure a peering relationship with each MSDP peer that resides either within the current PIM domain or in another PIM domain. MSDP is enabled on the router when you configure the first MSDP peering relationship.

Before you begin

Ensure that you configured PIM in the domains of the routers that you will configure as MSDP peers.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip msdp peer** *peer-ip-address* **connect-source** *interface* [**remote-as** *as-number*]
	Example:
	`switch(config)# ip msdp peer 192.168.1.10 connect-source ethernet 2/1`
	`remote-as 8`
		Configures an MSDP peer with the specified peer IP address. The software uses the source IP address of the interface for the TCP connection with the peer. The interface can take the form of *type slot/port*. If the AS number is the same as the local AS, then the peer is within the PIM domain; otherwise, this peer is external to the PIM domain. By default, MSDP peering is disabled.
		Note MSDP peering is enabled when you use this command.

Step 3	Repeat Step 2 for each MSDP peering relationship by changing the peer IP address, the interface, and the AS number as appropriate.	—
Step 4	**show ip msdp summary** [**vrf** *vrf-name* | *known-vrf-name* | **all**] **Example:** `switch# show ip msdp summary`	(Optional) Displays a summary of MDSP peers.
Step 5	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring MSDP Peer Parameters

You can configure the optional MSDP peer parameters described in this table. You configure these parameters in global configuration mode for each peer based on its IP address.

Table 21: MSDP Peer Parameters

Parameter	Description
Description	Description string for the peer. By default, the peer has no description.

Parameter	Description
Administrative shutdown	Method to shut down the MSDP peer. The configuration settings are not affected by this command. You can use this parameter to allow configuration of multiple parameters to occur before making the peer active. The TCP connection with other peers is terminated by the shutdown. By default, a peer is enabled when it is defined.
MD5 password	MD5-shared password key used for authenticating the peer. By default, no MD5 password is enabled.
SA policy IN	Route-map policy[5] for incoming SA messages. By default, all SA messages are received.
SA policy OUT	Route-map policy[6] for outgoing SA messages. By default, all registered sources are sent in SA messages.
SA limit	Number of (S, G) entries accepted from the peer and stored in the SA cache. By default, there is no limit.

[5] To configure route-map policies, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

[6] To configure route-map policies, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*.

Note Only the **match ip multicast group** command is supported for MSDP SA policy. The **match ip address** command for matching an ACL is not supported.

Procedure

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td><b>Step 1</b></td><td><b>config t</b>
<b>Example:</b>
<code>switch# config t</code>
<code>switch(config)#</code></td><td>Enters global configuration mode.</td></tr>
<tr><td><b>Step 2</b></td><td><table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td><b>ip msdp description</b> <i>peer-ip-address</i>
Example:
<code>switch(config)# ip msdp description 192.168.1.10 peer in Engineering network</code></td><td>Sets a description string for the peer. By default, the peer has no description.</td></tr>
</table></td><td>The following commands configure the MSDP peer parameters.</td></tr>
</table>

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td></td><td>
<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td><b>ip msdp shutdown</b> <i>peer-ip-address</i>
Example:
<code>switch(config)# ip msdp shutdown 192.168.1.10</code></td><td>Shuts down the peer. By default, the peer is enabled when it is defined.</td></tr>
<tr><td><b>ip msdp password</b> <i>peer-ip-address</i> <i>password</i>
Example:
<code>switch(config)# ip msdp password 192.168.1.10 my_md5_password</code></td><td>Enables an MD5 password for the peer. By default, no MD5 password is enabled.</td></tr>
<tr><td><b>ip msdp sa-policy</b> <i>peer-ip-address</i> <i>policy-name</i> <b>in</b>
Example:
<code>switch(config)# ip msdp sa-policy 192.168.1.10 my_incoming_sa_policy in</code></td><td>Enables a route-map policy for incoming SA messages. By default, all SA messages are received.</td></tr>
<tr><td><b>ip msdp sa-policy</b> <i>peer-ip-address</i> <i>policy-name</i> <b>out</b>
Example:
<code>switch(config)# ip msdp sa-policy 192.168.1.10 my_outgoing_sa_policy out</code></td><td>Enables a route-map policy for outgoing SA messages. By default, all registered sources are sent in SA messages.</td></tr>
<tr><td><b>ip msdp sa-limit</b> <i>peer-ip-address</i> <i>limit</i> <b>out</b>
Example:
<code>switch(config)# ip msdp sa-limit 192.168.1.10 5000</code></td><td>Sets a limit on the number of (S, G) entries accepted from the peer. By default, there is no limit.</td></tr>
</table>
</td><td></td></tr>
</table>

	Command or Action	Purpose
Step 3	**show ip msdp peer** [*peer-address*] [**vrf** *vrf-name* | *known-vrf-name* | **all**] **Example:** `switch# show ip msdp peer 1.1.1.1`	(Optional) Displays detailed MDSP peer information.
Step 4	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring MSDP Global Parameters

You can configure the optional MSDP global parameters described in this table.

Table 22: MSDP Global Parameters

Parameter	Description
Originator interface name	IP address used in the RP field of an SA message entry. When Anycast RPs are used, all RPs use the same IP address. You can use this parameter to define a unique IP address for the RP of each MSDP peer. By default, the software uses the RP address of the local system. **Note** We recommend that you use a loopback interface for the RP address.
Group limit	Maximum number of (S, G) entries that the software creates for the specified prefix. The software ignores groups when the group limit is exceeded and logs a violation. By default, no group limit is defined.
SA interval	Interval at which the software transmits Source-Active (SA) messages. The range is from 60 to 65,535 seconds. The default is 60 seconds.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

<table>
<tr><th></th><th>Command or Action</th><th>Purpose</th></tr>
<tr><td><b>Step 2</b></td><td>
<table>
<tr><th>Option</th><th>Description</th></tr>
<tr><td><b>ip msdp originator-id</b> <i>interface</i>
Example:
<code>switch(config)# ip msdp originator-id loopback0</code></td><td>Sets a description string for the peer. By default, the peer has no description.
Sets the IP address used in the RP field of an SA message entry. The interface can take the form of type slot/port. By default, the software uses the RP address of the local system.
<b>Note</b> We recommend that you use a loopback interface for the RP address.</td></tr>
<tr><td><b>ip msdp group-limit</b> <i>limit</i> <b>source</b> <i>source-prefix</i>
Example:
<code>switch(config)# ip msdp group-limit 1000 source 192.168.1.0/24</code></td><td>Maximum number of (S, G) entries that the software creates for the specified prefix. The software ignores groups when the group limit is exceeded and logs a violation. By default, no group limit is defined.</td></tr>
<tr><td><b>ip msdp sa-interval</b> <i>seconds</i>
Example:
<code>switch(config)# ip msdp sa-interval 80</code></td><td>Interval at which the software transmits Source-Active (SA) messages. The range is from 60 to 65,535 seconds. The default is 60 seconds.</td></tr>
</table>
</td><td></td></tr>
<tr><td><b>Step 3</b></td><td><b>show ip msdp summary</b> [<b>vrf</b> <i>vrf-name</i> | <i>known-vrf-name</i> | <b>all</b>]
<b>Example:</b>
<code>switch# show ip msdp summary</code></td><td>(Optional) Displays a summary of the MDSP configuration.</td></tr>
</table>

	Command or Action	Purpose
Step 4	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Configuring MSDP Mesh Groups

You can configure optional MDSP mesh groups in global configuration mode by specifying each peer in the mesh. You can configure multiple mesh groups on the same router and multiple peers per mesh group.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**ip msdp mesh-group** *peer-ip-addr mesh-name* **Example:** `switch(config)# ip msdp mesh-group 192.168.1.10 my_mesh_1`	Configures an MSDP mesh with the peer IP address specified. You can configure multiple meshes on the same router and multiple peers per mesh group. By default, no mesh groups are configured.
Step 3	Repeat Step 2 for each MSDP peer in the mesh by changing the peer IP address.	—
Step 4	**show ip msdp mesh-group** [*mesh-group*] [**vrf** *vrf-name* | *known-vrf-name* | **all**] **Example:** `switch# show ip msdp summary`	(Optional) Displays information about the MDSP mesh group configuration.
Step 5	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Restarting the MSDP Process

Before you begin

You can restart the MSDP process and optionally flush all routes.

Procedure

	Command or Action	Purpose
Step 1	**restart msdp** **Example:** `switch# restart msdp`	Restarts the MSDP process.

Step 2	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 3	**ip msdp flush-routes** **Example:** `switch(config)# ip msdp flush-routes`	Removes routes when the MSDP process is restarted. By default, routes are not flushed.
Step 4	**show running-configuration | include flush-routes** **Example:** `switch(config)# show running-configuration | include flush-routes`	(Optional) Shows flush-routes configuration lines in the running configuration.
Step 5	**copy running-config startup-config** **Example:** `switch(config)# copy running-config startup-config`	(Optional) Saves configuration changes.

Verifying the MSDP Configuration

To display the MSDP configuration, perform one of the following tasks:

Command	Description
show ip msdp count [*as-number*] [**vrf** *vrf-name* | *known-vrf-name* | **all**]	Displays MSDP (S, G) entry and group counts by the AS number.
show ip msdp mesh-group [*mesh-group*] [**vrf** *vrf-name* | *known-vrf-name* | **all**]	Displays the MSDP mesh group configuration.
show ip msdp peer [*peer-address*] [**vrf** *vrf-name* | *known-vrf-name* | **all**]	Displays MSDP information for the MSDP peer.
show ip msdp rpf [*rp-address*] [**vrf** *vrf-name* | *known-vrf-name* | **all**]	Displays next-hop AS on the BGP path to an RP address.
show ip msdp sources [**vrf** *vrf-name* | *known-vrf-name* | **all**]	Displays the MSDP-learned sources and violations of configured group limits.

Command	Description
show ip msdp summary [**vrf** *vrf-name* | *known-vrf-name* | **all**]	Displays a summary of the MSDP peer configuration.

For detailed information about the fields in the output from these commands, see *Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*.

Monitoring MSDP

You can display and clear MSDP statistics by using the features in this section.

Displaying Statistics

You can display MSDP statistics using the commands listed in this table.

Command	Description
show ip msdp [*as-number*] **internal event-history** {**errors** | **messages**}	Displays memory allocation statistics.
show ip msdp policy statistics sa-policy *peer-address* {**in** | **out**} [**vrf** *vrf-name* | *known-vrf-name* | **all**]	Displays the MSDP policy statistics for the MSDP peer.
show ip msdp {**sa-cache** | **route**} [*source-address*] [*group-address*] **vrf** *vrf-name* | *known-vrf-name* | **all**] [*asn-number*] [**peer***peer-address*]	Displays the MSDP SA route cache. If you specify the source address, all groups for that source are displayed. If you specify a group address, all sources for that group are displayed.

Clearing Statistics

You can clear the MSDP statistics using the commands listed in this table.

Table 23: MSDP Clear Statistics Commands

Command	Description
clear ip msdp peer [*peer-address*] [**vrf** *vrf-name* | *known-vrf-name*]	Clears the TCP connection to an MSDP peer.
clear ip msdp policy statistics sa-policy*peer-address* {**in** | **out** } [**vrf** *vrf-name* | *known-vrf-name*]	Clears statistics counters for MSDP peer SA policies.
clear ip msdp statistics*peer-address* [**vrf** *vrf-name* | *known-vrf-name*]	Clears statistics for MSDP peers.
clear ip msdp {**sa-cache** | **route**} [*group -address*] [**vrf** *vrf-name* | *known-vrf-name*] **all**]	Clears the group entries in the SA cache.

Configuration Examples for MSDP

To configure MSDP peers, some of the optional parameters, and a mesh group, follow these steps for each MSDP peer:

1. Configure the MSDP peering relationship with other routers.

   ```
   switch# config t
   switch(config)# switch(config)# ip msdp peer 192.168.1.10 connect-source ethernet 1/0
   remote-as 8
   ```

2. Configure the optional peer parameters.

   ```
   switch# config t
   switch(config)# ip msdp password 192.168.1.10 my_peer_password_AB
   ```

3. Configure the optional global parameters.

   ```
   switch# config t
   switch(config)# ip msdp sa-interval 80
   ```

4. Configure the peers in each mesh group.

   ```
   switch# config t
   switch(config)# ip msdp mesh-group 192.168.1.10 mesh_group_1
   ```

This example shows how to configure a subset of the MSDP peering.

RP 3: 192.168.3.10 (AS 7)

```
config t
  ip msdp peer 192.168.1.10 connect-source ethernet 1/1
  ip msdp peer 192.168.2.10 connect-source ethernet 1/2
  ip msdp peer 192.168.6.10 connect-source ethernet 1/3 remote-as
9
  ip msdp password 192.168.6.10 my_peer_password_36
  ip msdp sa-interval 80
  ip msdp mesh-group 192.168.1.10 mesh_group_123
  ip msdp mesh-group 192.168.2.10 mesh_group_123
  ip msdp mesh-group 192.168.3.10 mesh_group_123
```

RP 5: 192.168.5.10 (AS 8)

```
config t
  ip msdp peer 192.168.4.10 connect-source ethernet 1/1
  ip msdp peer 192.168.6.10 connect-source ethernet 1/2 remote-as
9
  ip msdp password 192.168.6.10 my_peer_password_56
  ip msdp sa-interval 80
```

RP 6: 192.168.6.10 (AS 9)

```
config t
  ip msdp peer 192.168.7.10 connect-source ethernet 1/1
  ip msdp peer 192.168.3.10 connect-source ethernet 1/2 remote-as
7
  ip msdp peer 192.168.5.10 connect-source ethernet 1/3 remote-as
8
  ip msdp password 192.168.3.10 my_peer_password_36
  ip msdp password 192.168.5.10 my_peer_password_56
  ip msdp sa-interval 80
```

Related Documents

Related Topic	**Document Title**
VDCs	*Cisco Nexus 7000 Series NX-OS Virtual Device Context Command Reference*
CLI Commands	*Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*
Configuring Policy Based Routing and MBGP	*Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide*

Standards

Standards	**Title**
RFC 4624	Multicast Source Discovery Protocol (MSDP) MIB

CHAPTER 9

Configuring Multicast Extranet

This chapter describes how to configure the Multicast Extranet feature on Cisco Nexus 7000 Series Switches.

Information About Configuring Multicast Extranet

An extranet can be viewed as part of a company’s intranet that is extended to users outside the company. It has also been described as a “state of mind” in which a VPN is used as a way to do business with other companies as well as to sell products and content to customers and companies. An extranet is a VPN connecting the corporate site or sites to external business partners or suppliers to securely share part of a business’s information or operations among them.

The Multicast VPN Extranet Support feature enables service providers to distribute IP multicast content originated from one enterprise site to other enterprise sites. This feature enables service providers to offer the next generation of flexible extranet services, helping to enable business partnerships between different enterprise VPN customers. Using this feature, service providers can offer multicast extranet contracts to meet various business partnership requirements, including short-term, annual, and rolling contracts.

Earlier to Release 8.2(1), Cisco NX-OS multicast implementation, multicast traffic can flow only within the same virtual routing and forwarding (VRF). However, with the introduction of the Multicast Extranet feature, multicast receivers can exist in different VRFs from source in an enterprise network.

With multicast extranet, the reverse path forwarding (RPF) lookup for multicast route in the receiver VRF can be carried out in the source VRF, thereby allowing to return a valid RPF interface. This forms a source or RP tree from the receiver VRF to the source VRF, thus enabling the traffic that originated from the source VRF to be forwarded to the outgoing interface (OIF) in the receiver VRF.

The Multicast VPN Extranet Support feature can be used to solve such business problems as:

- Efficient content distribution between enterprises
- Efficient content distribution from service providers or content providers to their different enterprise VPN customers

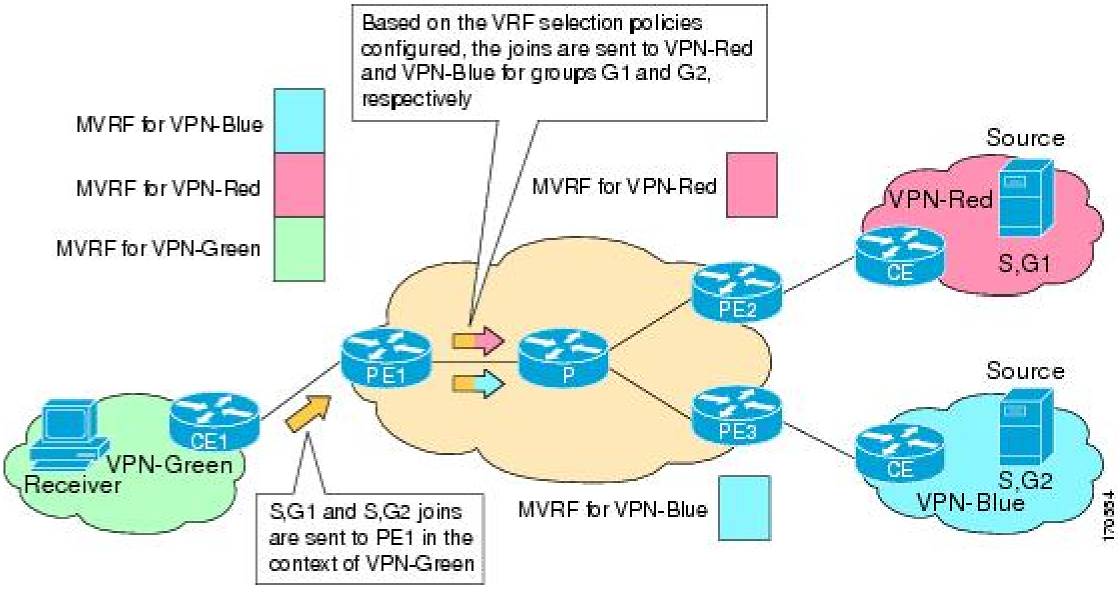
Prior to the introduction of the Multicast VPN Extranet VRF Select feature, RPF lookups for a source address could be performed only in a single VRF, that is, in the VRF where Internet Group Management Protocol (IGMP) or PIM joins are received, in the VRF learned from BGP imported routes, or in the VRF specified in static mroutes (when RPF for an extranet MVPN is configured using static mroutes). In those cases, the source VRF is solely determined by the source address or the way the source address was learned.

The Multicast Extranet VRF Select feature provides the capability for RPF lookups to be performed to the same source address in different VRFs using the group address as the VRF selector. This feature enhances extranet MVPNs by enabling service providers to distribute content streams coming in from different MVPNs and redistributing them from there.

The Multicast VPN VRF Select feature is configured by creating group-based VRF selection policies. Group-based VRF selection policies are configured using the ip multicast rpf select command. The ip multicast rpf select command is used to configure RPF lookups originating in a receiver MVRF or in the global routing table to be resolved in a source MVRF or in the global routing table based on group address. Access Control Lists (ACLs) are used to define the groups to be applied to group-based VRF selection policies.

The figure illustrates an extranet MVPN topology with the Multicast VPN VRF Select feature configured. In this topology, (S, G1) and (S, G2) PIM joins originating from VPN-Green, the receiver VRF, are forwarded to PE1, the receiver PE. Based on the group-based VRF selection policies configured, PE1 sends the PIM joins to VPN-Red and VPN-Blue for groups G1 and G2, respectively.

Figure 19: RPF Lookups Using Group-Based VRF Selection Policies


Based on the VRF selection policies configured, the joins are sent to VPN-Red and VPN-Blue for groups G1 and G2, respectively
MVRF for VPN-Blue
MVRF for VPN-Red
MVRF for VPN-Green
MVRF for VPN-Red
Source
VPN-Red
CE
S,G1
PE2
PE1
P
Source
PE3
CE1
VPN-Green
Receiver
CE
S,G2
VPN-Blue
S,G1 and S,G2 joins are sent to PE1 in the context of VPN-Green
MVRF for VPN-Blue
170534

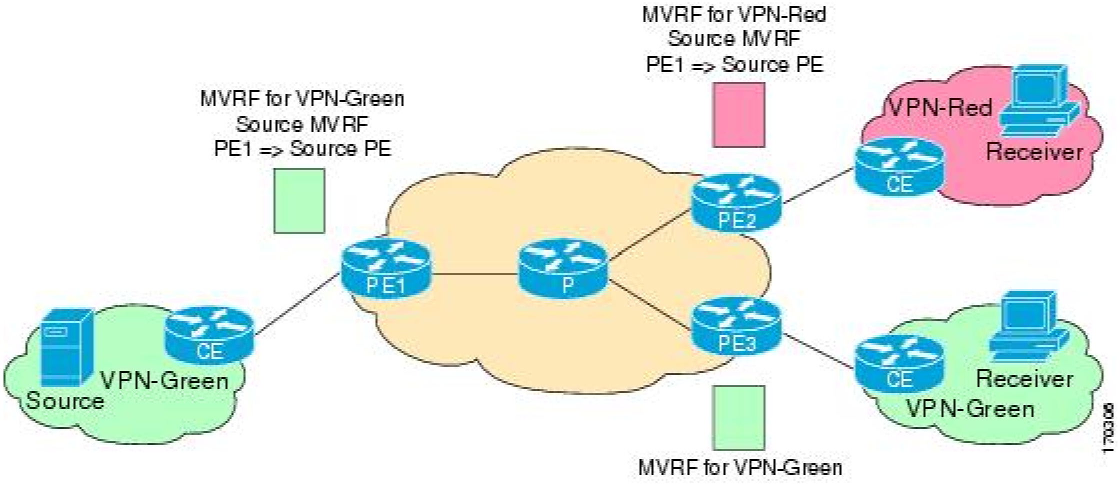
Components of Multicast Extranet

The figure below illustrates the components that constitute multicast extranet.

- **MVRF** --Multicast VPN routing and forwarding (VRF) instance. An MVRF is a multicast-enabled VRF. A VRF consists of an IP routing table, a derived forwarding table, a set of interfaces that use the forwarding table, and a set of rules and routing protocols that determine what goes into the forwarding table. In general, a VRF includes the routing information that defines a customer VPN site that is attached to a provider edge (PE) router.
- **Source MVRF** --An MVRF that can reach the source through a directly connected customer edge (CE) router.
- **Receiver MVRF** --An MVRF to which receivers are connected through one or more CE devices.
- **Source PE** --A PE router that has a multicast source behind a directly connected CE router.

- **Receiver PE** --A PE router that has one or more interested receivers behind a directly connected CE router.

Figure 20: Components of an Extranet MVPN



Guidelines and Limitations for Configuring Multicast Extranet

The Multicast Extranet feature has the following guidelines and limitations:

- RPF lookup will be performed on the VRF specified by the **ip multicast rpf select vrf** command. Fallback mode is not supported.
- The Multicast Extranet feature is supported only on M3 Series modules prior to Cisco NX-OS Release 8.4(1).
- Starting from Cisco NX-OS Release 8.4(1), the Multicast Extranet feature is supported on F4 Series modules.
- The number of multicast routes and VRFs that are required determine memory consumption by multicast.
- The source and (rendezvous point) RP should be in the same VRF.
- Static RP is supported for the multicast extranet group range.
- Multicast VPN (MVPN) extranet is not supported on multicast extranet.

How to Configure Multicast Extranet

Configuring Multicast Extranet

To configure multicast extranet, perform these steps.

Before you begin

Ensure that the Protocol Independent Multicast (PIM) sparse mode is enabled:

Procedure

Step 1 Enter configuration mode.

configure terminal

Step 2 Support RPF selection in a different VRF:

ip multicast rpf select vrf *src-vrf-name* **group-list** *group-range*

- **vrf** *src-vrf-name*—Specifies the source VRF name. The name can be a maximum of 32 alphanumeric characters, and is case sensitive.
- **group-list** *group-range*—Specifies the group range for the selected RPF. The format is A.B.C.D/LEN with a maximum length of 32.

To disable the support, use the **no** form of this command.

Step 3 View the running configuration information for the IPv4 multicast routes:

show ip mroute

Step 4 Save the configuration changes:

copy running-config startup-config

Configuration example for PVLAN over OTV

This example shows how to display information about running configuration for IPv4 multicast routes:

```
switch(config)# show ip mroute

IP Multicast Routing Table for VRF "default"

(*, 225.1.1.207/32), uptime: 00:13:33, ip pim
Incoming interface: Vlan147, RPF nbr: 147.147.147.2, uptime: 00:13:33
Outgoing interface list: (count: 0)

Extranet receiver in vrf blue:
(*, 225.1.1.207/32) OIF count: 1

(40.1.1.2/32, 225.1.1.207/32), uptime: 00:00:06, mrib ip pim
Incoming interface: Vlan147, RPF nbr: 147.147.147.2, uptime: 00:00:06
Outgoing interface list: (count: 0)

Extranet receiver in vrf blue:
(40.1.1.2/32, 225.1.1.207/32) OIF count: 1

switch(config)#
```

For detailed information about the fields in the output, see the *Cisco Nexus 7000 Series Command Reference*.

Additional References for Configuring Multicast Extranet

Related Documents

Related Topic	Document Title
CLI Commands	*Cisco Nexus 7000 Series Multicast Command Reference Guide*

Standards and RFCs

No new or modified standards are supported by this feature, and support for existing standards has not been modified by this feature.

Technical Assistance

Description	Link
The Cisco Support website provides extensive online resources, including documentation and tools for troubleshooting and resolving technical issues with Cisco products and technologies. To receive security and technical information about your products, you can subscribe to various services, such as the Product Alert Tool (accessed from Field Notices), the Cisco Technical Services Newsletter, and Really Simple Syndication (RSS) Feeds. Access to most tools on the Cisco Support website requires a Cisco.com user ID and password.	http://www.cisco.com/cisco/web/support/index.html

Feature Information for Configuring Multicast Extranet

The following table provides release information about the feature or features described in this module. This table lists only the software release that introduced support for a given feature in a given software release train. Unless noted otherwise, subsequent releases of that software release train also support that feature.

Use Cisco Feature Navigator to find information about platform support and Cisco software image support. To access Cisco Feature Navigator, go to www.cisco.com/go/cfn. An account on Cisco.com is not required.

Table 24: Feature Information for Configuring Multicast Extranet

Feature Name	Release	Feature Information
Multicast Extranet	8.4(1)	Added support for F4 Series modules.

Feature Name	Release	Feature Information
Configuring Multicast Extranet	8.2(1)	With multicast extranet, the reverse path forwarding (RPF) lookup for multicast route in the receiver VRF can be carried out in the source VRF, thereby allowing to return a valid RPF interface. The following command was introduced in this feature: **ip multicast rpf select vrf**.

CHAPTER 10

Configuring MoFRR

Configuring MoFRR

Information about MoFRR

Multicast only Fast Re-Route (MoFRR) is an IP solution that minimizes packet loss in a network when there is a link or node failure. It works by making simple enhancements to multicast routing protocols like Protocol Independent Multicast (PIM). It reduces multicast traffic disruption for the receivers in the event of Node or Link failure anywhere along the Multicast Tree.

MoFRR transmits a multicast join message from a receiver toward a source on a primary path, while also transmitting a secondary multicast join message from the receiver toward the source on a backup path. Data packets are received from both the primary path and the secondary paths. The redundant packets are discarded at topology merge points due to Reverse Path Forwarding (RPF) checks. When a failure is detected on the primary path, the repair is made by changing the interface on which packets are accepted to the secondary interface. Because the repair is local, it is fast—greatly improving convergence times in the event of node or link failures on the primary path.

The MoFRR feature provides the ability to minimize packet loss in a network when there is a link or node failure by enhancing, but not changing, multicast routing protocols such as PIM. With MoFRR, multicast routing protocols do not have to wait or depend on unicast routing protocols to detect network failures.

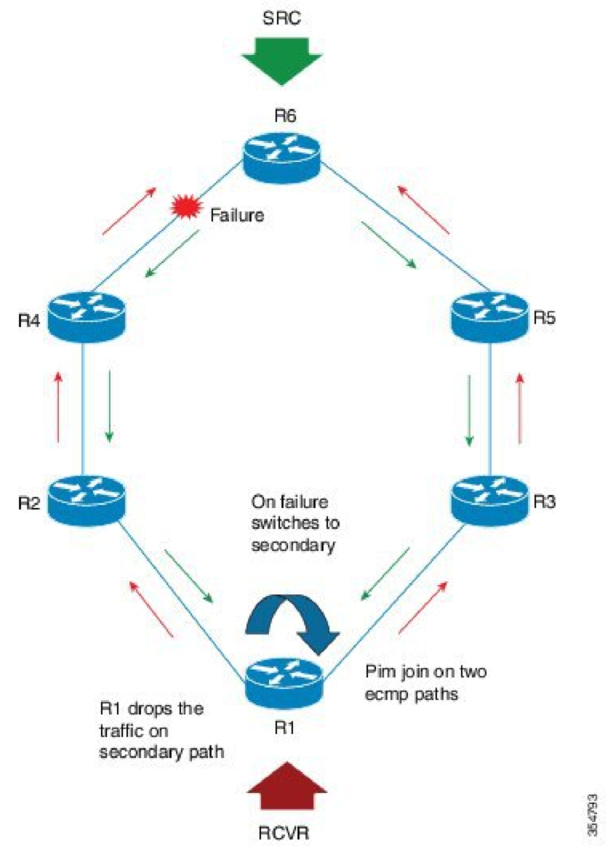
The MoFRR feature can be divided into two planes, red and blue, that are fully disjoint from each other all the way into the points of presence (POPs) as shown in the figure.

This two-plane design eliminates single points of failure in the core network. The upstream full-line arrows indicate the normal path taken when the PIM joins the flow from the POPs toward the source of the network.

MoFRR adds the broken-arrow path where the provider edge (PE) routers send an alternate PIM join to their neighbor toward the source. Each PE router then receives two copies of the same stream, one from the blue plane and one from the red plane. As a result of multicast RPF checks, the following occurs:

- The multicast stream received over the primary path (in the reverse direction of the full-line arrows) is accepted and forwarded to the downstream links.
- The copy of the stream received on the alternate path (in the reverse direction of the broken-line arrows) is discarded.

Figure 21: Multicast only Fast Re-Route



In the example above, when a routing failure occurs due to a link failure between R4 and R6 routers, R3 becomes the primary upstream router to reach the source. This link to the router then becomes the RPF interface, and a copy of the multicast stream being received on the link is accepted and forwarded to the downstream links.

When a routing failure occurs, for example due to a link failure in the blue path, the red upstream router in the red plane becomes the primary upstream router to reach the source. This link to the router then becomes the RPF interface, and the copy of the multicast stream being received on the link is accepted and forwarded to the downstream links.

MoFRR achieves faster convergence by prebuilding the alternate multicast tree and receiving the traffic on that alternate path. The example discussed above is a simple case where there are two paths from each PE device toward the source, one along the blue plane and one along the red plane.

Beginning with Release 8.2(1), Cisco Nexus 7000 Series Switches targets to achieve sub-sec convergence delay for 2K (S, G) running on F3/M3 cards, using MoFRR feature. MoFRR feature allows faster programming and improved convergence. Beginning with Cisco NX-OS Release 8.4(1), MoFRR support has been extended to F4 Series modules.

Prerequisites for MoFRR

- Ensure IP Multicast is enabled. For more information on configuring IP Multicast, refer *Cisco Nexus 7000 Series NX-OS Multicast Routing Configuration Guide.*
- Ensure that you have disjoint ECMP paths towards the source.

Guidelines and Restrictions for MoFRR

- The MoFRR feature is disabled by default and must be enabled using the CLI.
- MoFRR feature is supported on Cisco Nexus 7000 Series Switches, F3, F4, and M3 modules only.
- The Equal Cost Multipath Protocol (ECMP) feature is a requirement for the MoFRR feature to function.
- If ECMP is not configured, the two paths that are chosen from the ECMP paths are based on the RPF neighbor address.
- MoFRR works only for Sparse Multicast (SM) S, G, and Source Specific Multicast (SSM) routes.
- MoFRR is applicable to only IPv4 Multicast, not IPv6 Multicast.
- MoFRR does not support extranet routes.
- MoFRR works where the Reverse Path Forwarding (RPF) lookups are done in a single VRF.
- Both primary and secondary paths should exist in the same multicast topology.
- MoFRR is supported on images supporting IPv4 MFIB only.
- We recommend that you enable MoFRR feature on the last hop router.
- **ip multicast multipath legacy** MoFRR is not supported.
- For better convergence numbers instead of using default values for MoFRR, use below BGP Optimization CLI and OSPF Aggressive Timers.

Note Ensure that the detailed analysis is done before using the below configurations to avoid negative impact for other features enabled on the DUT.

BGP Optimization CLI : **nexthop trigger-delay**{**critical | non-critical**} milliseconds

Example:

```
switch(config-router-af)# nexthop trigger-delay critical 5000
```

Specifies next-hop address tracking delay timer for critical next-hop reachability routes and noncritical routes. The range is from 1 to 4294967295 milliseconds. The critical timer default is 3000 milliseconds and noncritical timer default value is 10000 milliseconds. As critical default value is 3000 millisecond, the resolution can delay upto 3 seconds for NH calculations.

Suggestions: Have an aggressive trigger -delay. This must be tuned as per requirement.

<**nexthop trigger-delay critical 1 non-critical 1** >

OSPF Aggressive Timers: This must be tuned as per requirement. The OSPF timers can be tuned as below:

Example

```
router ospf 100
timers throttle spf 20 50 500
timers lsa-arrival 50
timers throttle lsa 20 50 500
```

Configuring MoFRR

Perform the following steps to enable MoFRR:

Procedure

Step 1 Enter the global configuration mode.

config terminal

Step 2 Enable the MoFRR feature.

ip multicast rpf mofrr [**damping-interval** *interval*] [**resilient**] [**route-map** *map-name*]

damping-interval is specified in seconds and the value can range between 10 and 180.

Use the **resilient** option to make MoFRR RPF resilient.

This option avoids reuse of flows after an ECMP path returns after the MoFRR switchover event. Configure resilient option to make RPF resilient with ECMP path changes. If the ECMP path list changes and old RPF information is still part of the ECMP, configuring the resilient option will utilize old RPF information instead of reusing flows and potentially changing the RPF information.

Use the option route-map *map-name* to specify the route map policy name.

Verifying Configuring MoFRR

Perform the following steps to verify the configuration of MoFRR:

Procedure

Step 1 **show ip mroute mofrr**

Example:

```
switch# show ip mroute mofrr
```

Displays the information that IP multicast routing uses and the MoFRR information.

Step 2 **show ip pim route**

Example:

```
switch# show ip pim route
```

Displays the PIM status and configuration.

Step 3 **show forwarding distribution multicast route group** *group-addr* **source** *source-addr*

Example:

```
switch# show forwarding distribution multicast route group 225.1.1.1/32 source 20.20.20.4/32


  (20.20.20.4/32, 225.1.1.1/32), RPF Interface: Ethernet2/9, flags:
    Received Packets: 3479 Bytes: 222656
    MoFRR ID: 0x9734694
    Number of Outgoing Interfaces: 1
    Outgoing Interface List Index: 1
      Ethernet2/12
```

Step 4 **show forwarding multicast route group** *group-addr* **source** *source-addr*

Example:

```
switch# show forwarding multicast route group 225.1.1.1/32 source 20.20.20.4/32

slot  2
=======

  (20.20.20.4/32, 225.1.1.1/32), RPF Interface: Ethernet2/9, flags:
    Received Packets: 3628 Bytes: 232192
    MoFRR ID: 0x9734694
   Number of Outgoing Interfaces: 1
    Outgoing Interface List Index: 1
      Ethernet2/12  Outgoing Packets:38007993 Bytes:2483189012
```

Troubleshooting

The command-line interface (CLI) allows you to configure and monitor Cisco NX-OS using a local console or remotely using a Telnet or SSH session. Using the CLI, you can enable debugging modes and view a real-time updated activity log. You can use **show** commands to list historical and real-time information.

- You can enable debugging by running the **debug ip mrouting mofrr** command.
- Run the **show routing multicast internal event-history mofrr** command to view MoFRR data.

Feature Information for Configuring MoFRR

This table lists the release history for this feature.

Table 25: Feature Information MoFRR

Feature Name	Releases	Feature Description
Configuring MoFRR	8.4(1)	Added support for F4 series modules.
Configuring MoFRR	8.2(1)	The MoFRR feature provides the ability to minimize packet loss in a network when there is a link or node failure by enhancing multicast routing protocols such as PIM.

CHAPTER 11

Enabling Multicast Performance Enhancement on VDCs

This chapter describes how to enable the multicast performance enhancement for Cisco Nexus 7000 Series M1-XL Ethernet modules that are allocated to virtual device contexts (VDCs) in Cisco NX-OS devices.

Information About Multicast Performance Enhancement

In Cisco NX-OS 6.2(2) and later releases, the multicast performance enhancement supports the optimized shim frame format in multicast-replicated frames to improve multicast performance. The enhancement is supported on both Cisco Nexus 7000 M1 and M3 Series Ethernet modules with an XL option (M1-XL / M3-XL) that are allocated as resources in virtual device contexts (VDCs).

Guidelines and Limitations for Enhanced Multicast Performance

Enhanced multicast performance can be enabled only on Cisco Nexus 7000 Series M1-XL Ethernet modules that are allocated to a virtual device context (VDC).

Enabling Multicast Performance Enhancement

Note If you are familiar with the Cisco IOS CLI, be aware that the Cisco NX-OS commands for this feature might differ from the Cisco IOS commands that you would use.

Before you begin

- You must create the VDC on which you want to enable the multicast performance enhancement. For information, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide.*
- You have the name for the VDC to be configured.

Procedure

	Command or Action	Purpose

Step 1	**configure terminal**
	Example:
	`Switch# configure terminal`
	`Switch(config)#`
		Enables global configuration mode.

Step 2	**vdc** *vdc-name*
	Example:
	`Switch(config)# vdc MyVDC`
	`Switch(config-vdc)#`
		Specifies a VDC and enters VDC configuration mode.

Step 3	**limit-resource module-type m1xl**
	Example:
	`Switch(config-vdc)# limit resource module-type m1xl`
	`This will cause all ports of unallowed types to be removed from this vdc. Continue? [yes] Y`
	`Switch(config-vdc)#`
		Limits the resources for the VDC being configured to Cisco Nexus 7000 Series Ethernet modules with an XL Option only.

Step 4	**switchto vdc** *vdc-name*
	Example:
	`Switch(config-vdc)# switchto vdc MyVDC`
	`Switch-MyVDC(config-vdc#)`
		Switches from the default VDC to the specified VDC.
		Note You must be a network-admin or network-operator to use the **switchto** vdc command.

Step 5	**hardware forwarding shim** **Example:** `Switch-MyVDC(config-vdc)# hardware forwarding shim`	Enables shim optimization in frame header for this VDC.
Step 6	**show vdc** *vdc-name* [*detail*] **Example:** `Switch-MyVDC(config-vdc)# show vdc MyVDC`	(Optional) Displays information about the specified VDC.
Step 7	**copy running-config startup-config** **Example:** `Switch-MyVDC(config-vdc)# copy running-config startup-config`	(Optional) Copies the running configuration to the startup configuration.

Related Documents for Multicast Performance Enhancement

Related Topic	Document Title
Multicast commands	*Cisco Nexus 7000 Series NX-OS Multicast Routing Command Reference*
VDCs	*Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide*
VDC commands	*Cisco Nexus 7000 Series NX-OS Virtual Device Context Command Reference*

Feature History for Multicast Performance Enhancement

This Table lists the release history for this feature.

Table 26: Feature History for Multicast Performance Enhancement

Feature Name	Releases	Feature Information
Multicast performance enhancement	6.2(2)	Enables enhanced multicast performance on Cisco Nexus 7000 Series Ethernet modules with an XL Option allocated to virtual device contexts (VDCs). The following command was introduced: **hardware forwarding shim**.

CHAPTER 12

Configuring Multicast Interoperation with N7K-F132-15 Modules

This appendix describes how multicasting interoperates in a chassis that contains both M Series and N7K-F132-15 modules.

Information About Multicast Interoperation

Beginning with Cisco NX-OS Release 5.1, you can add an N7K-F132-15 module, which is a Layer 2-only module, into the Cisco Nexus 7000 Series chassis. You can add this module to a chassis that already contains M-Series modules to provide multicasting in a chassis that contains both N7K-F132-15 and M-Series modules.

Multicast Interoperation with N7K-F132-15 and M-Series Modules

Note You must install an M-Series module in the Cisco Nexus 7000 Series chassis to run Layer 3 routing and multicasting with the N7K-F132-15 module because you must have interfaces from both the M-Series and the N7K-F132-15 modules in the same virtual device context (VDC). See the *Cisco Nexus 7000 Series Virtual Device Context Configuration Guide* for more information on VDCs.

Layer 3 routing and multicasting come up automatically when you have an M-Series module installed in the chassis with the N7K-F132-15 module. You can position a chassis with both N7K-F132-15 and M-Series modules at the boundary between the Layer 2 and Layer 3 networks.

You must configure a VLAN interface for each VLAN on the N7K-F132-15 module that you want to use the proxy-routing functionality in a chassis that contains both N7K-F132-15 and M-Series modules. See the *Cisco Nexus 7000 Series Interfaces Configuration Guide* for information on configuring VLAN interfaces.

By default, all of the physical interfaces on the M-Series modules in the VDC become proxy routing ports for the VLANs that are configured with VLAN interfaces on the Layer 2-only N7K-F132-15 module in the same VDC. The physical interfaces on the M-Series module can be administratively down and they still pass traffic as proxy routers.

Packets that enter an interface on the N7K-F132-15 module are automatically forwarded to one of the interfaces on the M-Series modules in the same VDC to be routed. The interface on the M-Series module also performs egress replication for Layer 3 multicast packets that enter an interface on the N7K-F132-15 module in the same VDC. See the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide* for additional information about the routing interoperation with the N7K-F132-15 module.

You can specify which interfaces on the M-Series modules in the VDC where you want to perform the egress replication of VLAN interfaces for Layer 3 multicast packets. For multicast egress replication, the system automatically rebalances all the VLAN interfaces among all the available M-Series proxy routing interfaces. You can specify automatic or manual rebalancing among the proxy multicast replicators. If you specify manual rebalancing, you trigger a rebalance by entering a command. This command is useful when you are inserting or removing modules.

Note When you configure manual egress multicast replication load balancing and enter the rebalancing command, that command is not part of the configuration. It is not included in the commands that are copied when you enter the **copy running-config startup-config** command.

Virtualization Support

You must have interfaces from both the M-Series and the N7K-F132-15 modules in the same VDC.

See the *Cisco Nexus 7000 Series Virtual Device Context Configuration Guide* for more information about VDCs.

High Availability

For information about high availability, see the *Cisco Nexus 7000 Series NX-OS High Availability and Redundancy Guide*.

Prerequisites for Multicast Interoperation

For multicast interoperation, you must have at least one module of the following series in the Cisco Nexus 7000 Series chassis, as well as a valid license installed:

- M Series
- N7K-F132-15

Guidelines and Limitations

Multicasting requires you to have interfaces from both the M-Series and the N7K-F132-15 modules in the same VDC.

Configuring Layer 3 Multicast Using a Mixed Chassis

You can configure a Layer 3 gateway in a chassis with N7K-F132-15 and M-Series modules, by using the proxy routing functionality. You enable routing on a specific VLAN by configuring a VLAN interface. See the *Cisco Nexus 7000 Series NX-OS Interfaces Configuration Guide* for more information about Layer 3 routing and VLAN interfaces.

By default, Layer 3 routing and multicasting come up automatically when you have an M-Series module installed in the chassis with the N7K-F132-15 module. Layer 3 routing, multicasting, and load balancing among the available M-Series works by default using proxy routing on the M-Series modules.

Optionally, you can specify which physical interfaces on the M-Series modules that you want to use for egress multicast replication, as well as forcing rebalancing.

Before you begin

You must configure a VLAN interface for each VLAN on the N7K-F132-15 module where you want to use the proxy-routing functionality in a mixed chassis.

You must have interfaces from both the M-Series and the N7K-F132-15 modules in the same VDC.

If you remove an interface from the VDC and then enter this command, the removed interface only display when you reload the VDC.

Procedure

	Command or Action	Purpose

Step 1	**config t**
	Example:
	`switch# config t`
	`switch(config)#`
		Enters global configuration mode.

Step 2	**hardware proxy layer-3 replication** {**use** | **exclude**} {**module** *mod-number* | **interface** *slot/port*} [**module-type** *f1*]
	Example:
	`switch(config)# hardware proxy layer-3 replication exclude interface ethernet`
	`2/1-16, ethernet 3/1, ethernet 4/1-2`
		Configures specific modules and physical interfaces on the M-Series module to provide egress proxy replication of Layer 3 multicast packets on the N7K-F132-15 module.

Step 3	**hardware proxy layer-3 replication rebalance-mode** { **auto** | **manual**} **Example:**	Configures the load balancing among the proxy routing replication interfaces. When you choose **auto**, the switch automatically rebalances the configured VLAN interface multicast

	Command or Action	Purpose
	`switch(config)# hardware proxy layer-3 replication rebalance-mode auto`	replication traffic load among all the M-Series replicators. The default value is manual. **Note** In the *manual* mode, the switch automatically balances all traffic among the available proxy routing interfaces on the M-Series modules in the chassis when you first start up the switch.
Step 4	**hardware proxy layer-3 replication trigger rebalance** **Example:** `switch(config)# hardware proxy layer-3 replication trigger rebalance`	When you configure **manual** again in Step 3, use this command to trigger one-time load balancing among all the proxy routing multicast replication interfaces. This command is not effective if you have configured auto in Step 3. **Note** This command is not saved in the configuration; it is a one-time event.

Step 5	**exit**
	Example:
	`switch(config)# exit`
	`switch#`
		Exits configuration mode.

Step 6	**show hardware proxy layer-3 detail** **Example:** `switch# show hardware proxy layer-3 detail`	(Optional) Displays the information on the proxy Layer-3 functionality.
Step 7	**copy running-config startup-config** **Example:** `switch# copy running-config startup-config`	(Optional) Copies the running configuration to the startup configuration.

Example

This example shows how to specify certain physical interfaces on the N7K-M Series modules to perform egress multicast replication for VLANs on the N7K-F132-15 module in a mixed chassis and to trigger a rebalance:

```
switch# config t
switch(config)# hardware proxy layer-3 replication exclude interface ethernet 2/1-16, 3/1,
 4/1-2
switch(config)# hardware proxy layer-3 replication rebalance mode manual
switch(config)# hardware proxy layer-3 replication trigger rebalance
switch(config)#
```

Verifying the Multicast Configuration

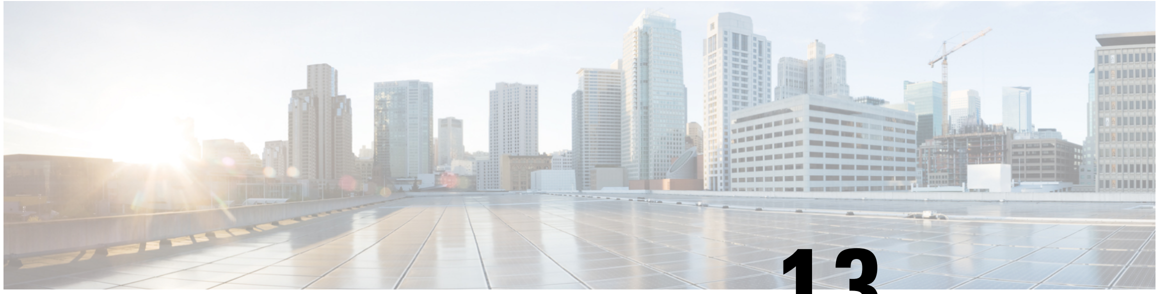
To display multicast configuration information, perform one of the following tasks:

Command	Description
show hardware proxy layer-3 detail	Displays information about the Layer 3 proxy routing functionality in a mixed chassis with both M-series and N7K-F132-15 modules.
show hardware proxy layer-3 counters {**brief** | **detail**}	Displays information about the number of packets that are sent by the N7K-F132-15 modules to each of the M-Series modules for proxy forwarding. **Note** Enter the **clear hardware proxy layer-3 counters** command to reset the counters to 0.

Feature History for Multicast Interoperation

Table 27: Feature History for Multicast Interoperation

Feature Name	Releases	Feature Information
Multicast interoperation between F-Series and M-Series modules in the Cisco Nexus 7000 Series chassis	5.1(1)	This feature, along with the N7K-F132-15 modules, was introduced in this release.



CHAPTER 13

L2 Multicast Route Update Optimisation

Information About L2Multicast Route Update Optimization

Beginning with Cisco NX-OS Release 8.2(9) Multicast Route Update Optimization feature is introduced to avoid packet loss in the multicast group.

Before release 8.2(9), when two or more host connects to a multicast group, if one of the hosts disconnects from the multicast group, there is packet loss for the existing host in the same multicast group if multicast source and receiver are in the same vlan. This packet loss is on l2 multicast on M3 and F4 Series modules.

Note By default, this configuration is disabled. You must configure this feature using admin VDC or default VDC role to enable on multicast group.

Guidelines and Limitations

The following are the guidelines and limitations for L2Multicast Route Update Optimization:

- This feature is supported only for M3 Series and F4 Series modules.
- You can configure this feature for admin VDC or default VDC.
- Configuring multicast route update optimization is not supported over a fabric path.

Configuring Multicast Route Update Optimization

To avoid packet loss for the existing host in a multicast group, configure multicast route update optimization using the below command:

Procedure

Step 1 **l2mcast route update optimize**

Example:

```
switch(config)# l2mcast route update optimize
```

Enables route update optimization, to avoid packet loss in a multicast group.

Step 2 **show running-config | include l2mcast**

(Optional) Shows the running-config information for l2 multicast route update optimization.

Example:

```
switch(config)# l2mcast route update optimize
```

APPENDIX A

IETF RFCs for IP Multicast

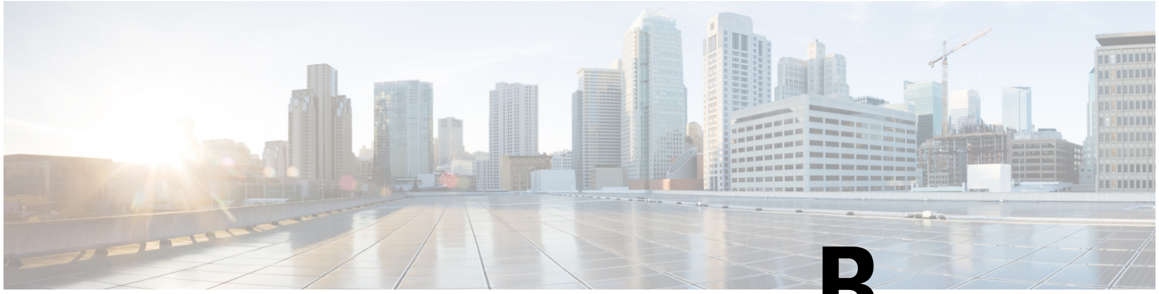
This appendix contains Internet Engineering Task Force (IETF) RFCs related to IP multicast. For information about IETF RFCs, see http://www.ietf.org/rfc.html.

IETF RFCs for IP Multicast

This appendix contains Internet Engineering Task Force (IETF) RFCs related to IP multicast. For information about IETF RFCs, see http://www.ietf.org/rfc.html.

RFCs	Title
RFC 2236	Internet Group Management Protocol
RFC 2365	*Administratively Scoped IP Multicast*
RFC 2858	*Multiprotocol Extensions for BGP-4*
RFC 3376	*Internet Group Management Protocoll*
RFC 3446	*Anycast Rendezvous Point (RP) mechanism using Protocol Independent Multicast (PIM) and Multicast Source Discovery Protocol (MSDP)*
RFC 3569	*An Overview of Source-Specific Multicast (SSM)*
RFC 3618	*Multicast Source Discovery Protocol (MSDP)*
RFC 4291	*IP Version 6 Addressing Architecture*
RFC 4541	*Considerations for Internet Group Management Protocol (IGMP) and Multicast Listener Discovery (MLD) Snooping Switches*
RFC 4601	*Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)*
RFC 4610	*Anycast-RP Using Protocol Independent Multicast (PIM)*

RFCs	Title
RFC 5059	*Bootstrap Router (BSR) Mechanism for Protocol Independent Multicast (PIM)*
RFC 5132	*IP Multicast MIB*



APPENDIX B

Configuration Limits for Cisco NX-OS Multicast

This appendix describes the configuration limits for Cisco NX-OS Multicast.

Configuration Limits

The features supported by Cisco NX-OS have maximum configuration limits. Some of the features have configurations that support limits less than the maximum limits.

The configuration limits are documented in the *Cisco Nexus 7000 Series NX-OS Verified Scalability Guide*.